Diagrammatic Analysis of J.S. Bach's *The Well-Tempered Clavier* Fugues, BWV 846–851

Anna Matuszewska [1,2]
[1] The Institute of Literary Research
of the Polish Academy of Sciences, Poland
an.ma.matuszewska@gmail.com

Christoph Seibert [2]
[2] Institute for Music Informatics and Musicology
University of Music, Karlsruhe, Germany
christoph.seibert@hfm-karlsruhe.de

Abstract

The field of musicology is constantly being enriched with digital, searchable music data. This trend opens new research possibilities; conversely, it requires new abilities to work with numerous data sets efficiently. Digital tools facilitate searching large music corpora and serve music analysis well. Nevertheless, there is still a potential to better harmonize research perspectives from musicology and computer science to make computational analysis outcomes more explicit, comprehensible, and flexible.

The aim of this paper is to present new ways of handling, displaying, and considering musicological data. Music information from fugues BWV 846–851 composed by J.S. Bach, retrieved with Humdrum Tools and the Music Processing Suite (MPS) software, was processed and translated into a relational database.[1] The visual display of the retrieved information was accomplished with dashboards using the data visualization software Tableau Public.[2] The possibility of comparing each fugue's voices makes it easier to comprehend the knowledge hidden behind music data. Additional options enable further visual exploration of the analyses and ensure conditions for abduction under assumptions of diagrammatic reasoning as proposed by Charles Sanders Peirce.

Introduction

Over the past decade, digital scholarly editions of music collections have flourished around the globe, and this trend continues. Digital tools for music analysis can visualize music features such as interval distribution, rhythmic patterns, and other statistics in ready-to-use charts. As a result, the analysis of large corpora has been simplified, but establishing relations between both subsets of a corpus and particular pieces of music can still be perplexing. Therefore, the visual exploration of analytical data should be regarded as a critical part of the research process.

Symbolic music data can be presented on different kinds of diagrams, thus becoming part of the epistemic process. Sybille Krämer argues that hybrid forms of word and image trigger our creativity, and it is one of the reasons why thinking with diagrams as hybrid representations "form the fundamental semiotic basis of human cognition" [7, p. 15]. According to Valeria Giardino, diagrams are computationally efficient because they externalize information in space, and therefore the user may act on them. They are very good inferential shortcuts in problem-solving. By spatially displaying their content, they affect memory and reasoning. Thanks to their nature of images, they can target issues concerning the relationship between perception, cognition, truth, and knowledge. Diagrams are not only spatial tools; they are 'scaffoldings' enhancing the ability to reason and infer. The space of the diagrams is a space interpreted along with textual and background knowledge [4].

The outcomes of the music analysis of Bach's fugues, BWV 846–851, were visualized on dashboards with interactive charts published on the platform Tableau Public. Chart combinations are filtered jointly. The output may be narrowed down to any number of voices, a range of measures, or a specified musical form element such as subject, countersubject, and episode. Explanatory descriptions support visual information presented on dashboards, and it contains supplementary textual information encoded in tooltips that may be displayed by hovering over an element.

1 Link to the project: https://tabsoft.co/3qfeYnt (accessed January 12, 2022).

2 https://public.tableau.com/en-us/s/ (accessed January 12, 2022).

The concept of diagrammatics inspires this project. Its main goal is to afford inference by the spatial display of symbolic music data. The encoded information is presented in diverse representational formats on a group of combined interactive charts. This approach tends to exploit their illustrative capabilities amplifying dynamic analysis. The current implementation of the project is adapted for two stages: a detailed analysis of a single piece of music and a comparative analysis of music collections on the macro level.

1 Diagrammatics as an Approach to Music Analysis

The concept of diagrammatics refers to the heritage of the American pragmatist and semiotician Charles S. Peirce. It addresses the design and processing of knowledge in diagrams and can be understood in two ways. First, the diagrammatic project may be developed from the manifestations gained in concrete diagrams (bottom-up analysis). Second, however, the consideration of concrete diagrams may also be derived from the theoretical premises of a general diagram (top-down analysis). The first approach is a theory of properties linked to diagrammatic structures, while the latter puts in the spotlight an epistemological process primarily associated with mentally realized diagrammatic conclusions [1, p. 16–17]. The diagrammatic reasoning proposed by Peirce is in harmony with the latter approach [1, p. 21].

Peirce stated that conclusive thinking is performed diagrammatically and interlocked diagrammatic reasoning with abduction [1, p. 64]. The abduction theory evolved from the early syllogistic theory where deduction, induction and abduction were treated as three different classes of reasoning [8] to the inferential theory where they were seen as three stadiums of one research method, in which abduction triggered the whole process [13, p. 15–25].

> Abduction is the process of forming an explanatory hypothesis. [...] Deduction proves that something *must* be [emphasis changed to italics]; Induction shows that something *actually is* operative [emphasis changed to italics]; Abduction merely suggests that something *may be* [emphasis changed to italics]. Its only justification is that from its suggestion deduction can draw a prediction which can be tested by induction, and that, if we are ever to learn anything or to understand phenomena at all, it must be by abduction that this is to be brought about [9].

The concept of diagrammatics is based on the belief that the iconic character class of the diagram is a metaphorical pattern (and in so far, a diagram itself) of inferential thought [1, p. 64]. Diagrams convey abstract information by depicting a physical situation from which the abstractions can be inferred [2, p. 5–22].

> To sum up, the overall picture of the diagrammatic reasoning process is that it forms a formal deductive reasoning core, embedded, on each side, in the trial-and-error of abductive trials and inductive tests [12, p. 105].

Abduction is inextricably linked with diagrammatic thinking, and this project aims to construct a combination of charts enabling the user to interact with them in accordance with this concept. Basic music information anchored on proposed interactive charts can be viewed broadly or within a narrowly defined scope. It can be reshaped and enriched with available additional information displayed both visually and numerically on linked charts. As a collection of charts, each dashboard comprises different representational formats presenting chosen elements of musical analysis. The process of abduction can start with any chart that draws the attention or curiosity of the viewer. The interaction with data, the possibility to access it differently, freely combine, and filter it, empowers creative reasoning. The information from each chart complements each other, providing the core of deductive reasoning. Accessing the data from different angles and multiple changes in its configuration make it possible to test an explanatory hypothesis inductively.

2 Data Preparation and Data Analysis

A computational analysis of six fugues from *The Well-Tempered Clavier* by J.S. Bach (BWV 846–851) was conducted using Humdrum Tools[3] and the Music Processing Suite software [5, 6]. The choice of tools was dictated by the fact that none of them requires any knowledge of programming languages; mastering them is possible for scholars not advanced in technology and both are available as free software. In addition, the flexibility of Humdrum **kern data and the ability to attribute information to each note is valuable. Generating data in Linux command line is not troublesome, and the data representation format is intuitive. Music Processing Suite does not require any significant technical preparation; after the installation, an analysis report may be generated with a set of built-in charts describing various musical features. The data generated in `.csv` format can be accessed and transformed easily.

All the statistical calculations are based on musical scores from the KernScores library [10] encoded in the Humdrum **kern data format. Humdrum Tools were used to calculate a number of notes and rests, the highest and the lowest note, the shortest and the longest note duration (command: `census -k`), melodic and rhythmic progressions (command: `extractx`), statistics about melodic intervals in each piece (command: `mint -s`) and dissonant types (command: `dissonant -u`). In addition, Music Processing Suite software was applied for harmonic analysis. Because coordination between Tableau Public, music score, and recording is not possible, there is an integrated link to Verovio Humdrum Viewer [11] in the navigation bar where the selected sheet of music can be simultaneously listened to.

3 Visual Representation and Interactive Elements for Visual Exploration

The main assumption of our project was to examine possibilities of displaying analysis results, to enable interaction with data, and ensure conditions for diagrammatic reasoning, not a thorough analysis itself. The presented project is in the development phase, and the adequacy of chart types for each of the music element representations are to be evaluated by the target audience. This version should indicate what impact on research practice is possible with the proposed data representation. The intent is to test the influence of diagrammatic reasoning on the research practice, not the choice of single diagrammatic elements.

The project is designed on three dashboards: two dashboards with the focus on the melodic aspect (Figure 1) or the rhythmic aspect (Figure 2), in which one fugue from a collection is displayed, and a complementary dashboard (Figure 3) on which multiple chosen pieces can be compared on a macro level. Tableau operates on data from Excel worksheets. Aspects of a musical piece are stored in separate Excel files (e.g., melodic sequences, harmony, dissonance, intervals). Each file comprises worksheets for each fugue, which are merged in Tableau using the union function. The Excel files are joined using mutual columns of data (number of fugues, voice, onset).

All the charts on each dashboard are interlinked, and the display changes dynamically according to the filters applied. Furthermore, each chart is supplied with an information button icon with a detailed description of the axes, the aim of the chart and the tool for data extraction is named.

3.1 Main Dashboard

Navigation through the main dashboard provides the possibility to narrow down the scope of the analysis to voices, measures, particular segments such as subjects or countersubjects, their unrestricted combination or to a single note. In addition to filtering, music information may be highlighted according to its attribute. It also displays cadences and dissonances.

Subject and Countersubject Analysis. A Gantt chart placed on the left top side of the main dashboard shows the occurrences of subjects and countersubjects over time in different voices in parallel. It contains all the appearances of the subject, countersubject, their variations, and cadences spread over time (x-axis). These may be a basis for data filtration. Voices (y-axis) are defined as in Humdrum music representation, where they are represented by spines and counted from the lower one upwards. The depicted analysis of fugues BWV 846–851 is

3 https://www.humdrum.org/tool/ (accessed January 12, 2022).

based on algorithmic calculations.[4] The output was compared with a traditional analysis conducted by Siglind Bruhn [3] to eliminate false-positive results.

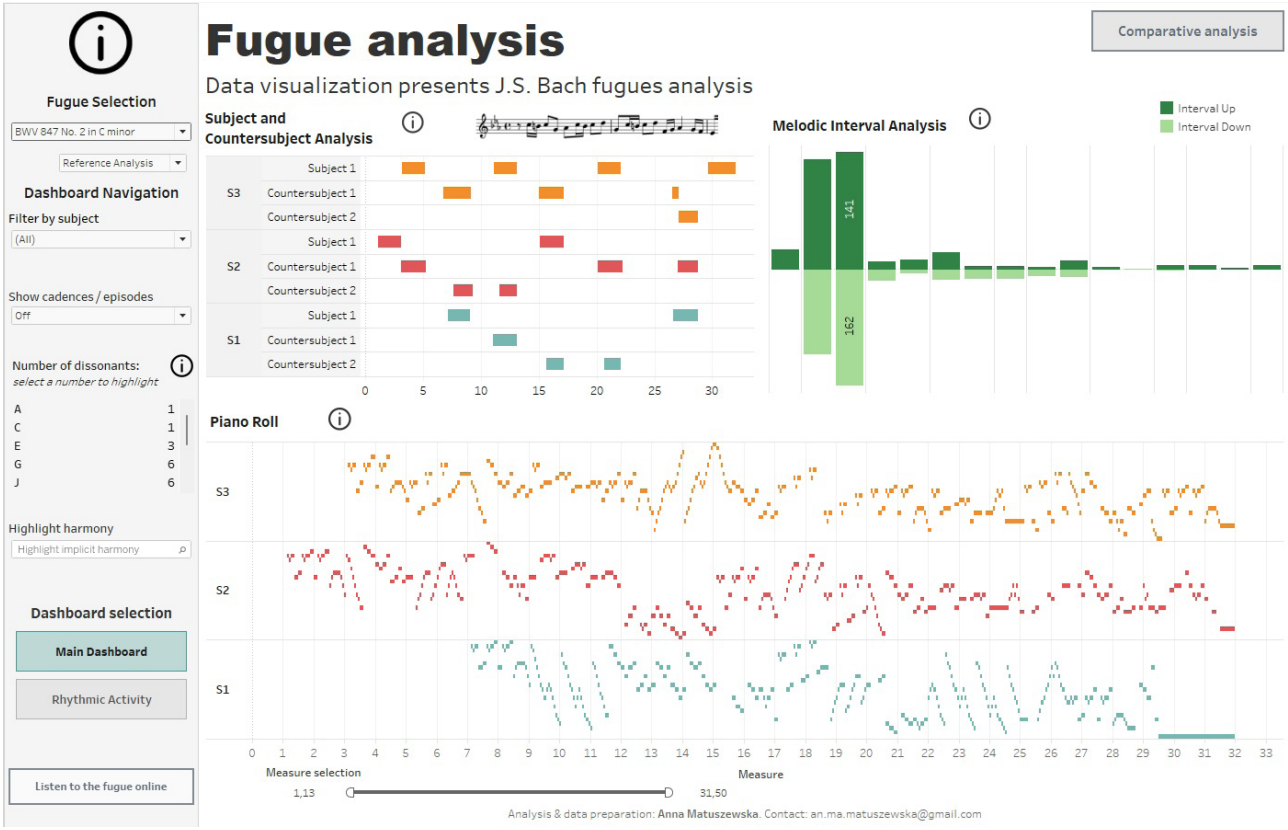


Figure 1: A view of the main dashboard in the Fugue Analysis project with data visualization of J.S. Bach's fugue No. 2, C minor, BWV 847.

Interval Analysis. On the top right side of the dashboard, a column chart calculates the number of melodic intervals in a chosen voice, voice combination, or fragment. On the y-axis, the direction of the interval is marked with a color. Additional information available by hovering over a column contains the percentage of the respective interval in a chosen voice and a complete composition. By clicking on a specified interval, its occurrences are showed on corresponding charts.

Piano Roll. Another Gantt chart on the bottom of the main dashboard with x-axis representing onset and y-axis representing voices. It allows for tracking the melodic progression along with its pitch, notes duration, context harmony. Additionally, chords or dissonance types can be highlighted using the lists displayed in the Dashboard Navigation section. A description of dissonances is provided in the legend.

The following passage describes a possible application of this dashboard for music analysis using the example of fugue BWV 847 (Figure 1):

> *After looking at the* Subject and Countersubject Analysis *chart we learn that there are six episodes. By observing the number of melodic intervals separately for subjects and episodes, we see that the proportion of intervals is different; we hypothesize that episodes' material is not derived from the subjects' material. From the theory, we know that there is a need to test rhythmic similarity as well. The* Rhythmic Activity *dashboard (see 3.2) suggests that rhythmic patterns in subject and episodes are similar. We go back to* Piano Roll*, zoom into the episode bars separately, and learn that although the jump in the head motif of the subject was changed, episodes 1, 2, 4 and 5 are bonded with subjects' head motif, episodes 3 and 6 are built on another material.*

4 Algorithmic detection of subjects and countersubjects is part of a broader-scope project that is not finished yet and will not be discussed in this paper.

3.2 Rhythmic Activity Dashboard

This dashboard contains three interactive charts devoted to rhythm analysis. The *Rhythmic Activity* chart constitutes a central place, and it may be a starting point for abductive reasoning. Its explicit visual outlay suggests possible fragments of the piece of music that might be worth leaning over. *Note Duration Distribution* and *Note Duration Distribution per Beat* charts may trigger the generation of abductive hypotheses. Deductive steps should be based on the analysis of all proposed charts.

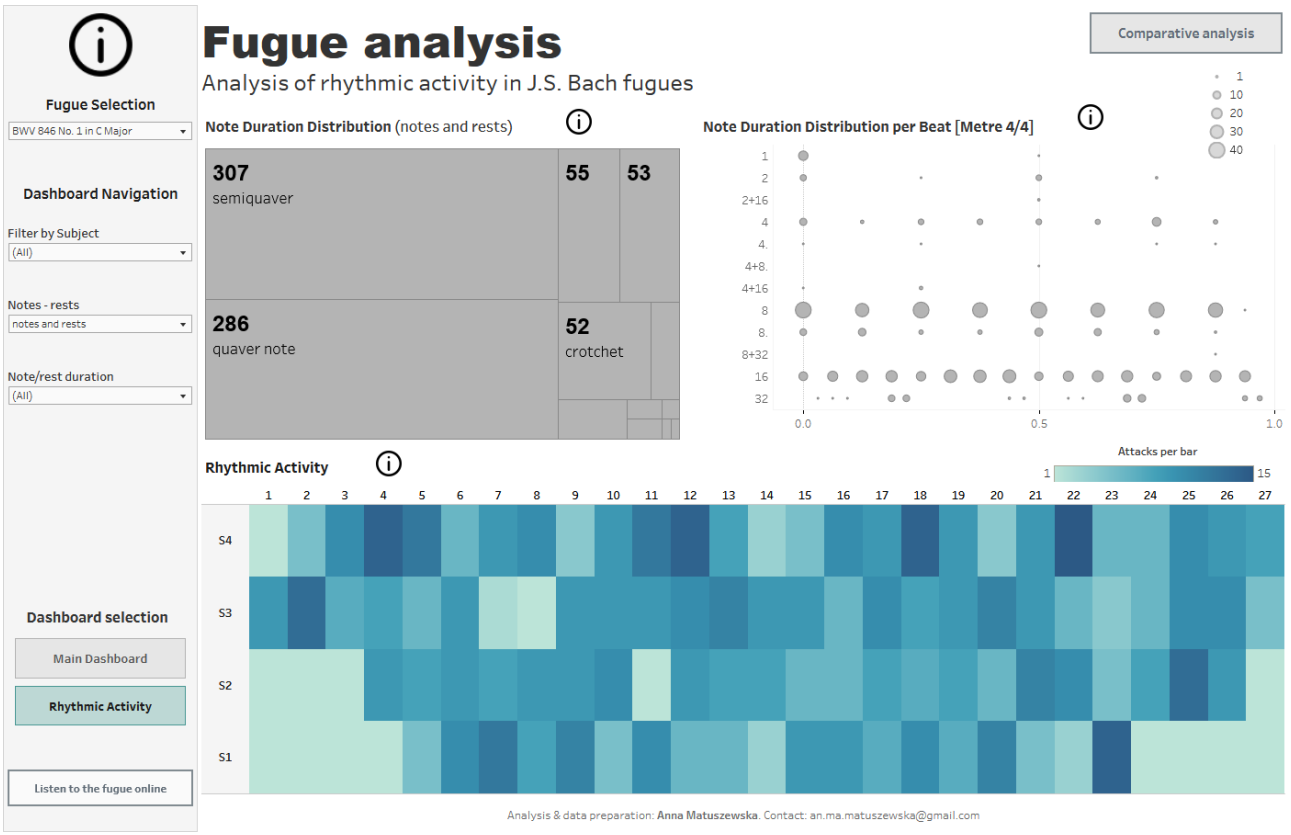


Figure 2: A view of the *Rhythmic Activity* dashboard in the Fugue Analysis project with data visualization of J.S. Bach's fugue No. 1, C major, BWV 846.

Rhythmic Activity is designed as a heat map. The x-axis depicts measures, y-axis voices. By hovering over a field on the chart, the number of attacks in a measure is displayed. The darker a colour of a field, the more attacks are in that measure.

Note Duration Distribution is a tree chart that gives information about the number of occurrences of each rhythmic value in a chosen piece of music or one of its parts (when filtered). The information is displayed as a number but also as the size of dedicated fields.

Note Duration Distribution per Beat is a bubble plot in which the x-axis is a mathematical division of one measure, and the y-axis presents the number of occurrences of each note duration. Thus, the chart indicates what rhythmical structures prevail on each beat. An additional tooltip shows specified information about the exact number of occurrences of a rhythmic class and its percentage in a fugue.

The following passage describes a possible application of this dashboard for music analysis using the example of fugue BWV 846 (Figure 2):

> *We begin the exploration of this dashboard by looking at the* Note Duration Distribution *chart. It informs us about the proportion of each note duration in the fugue. Surprisingly, the number of occurrences of dotted quavers, demisemiquavers and crotchets is quite similar. The hypothesis is that there is a rhythmic pattern built on these three note durations. From the* Note Duration Distribution per Beat*, we infer a bond between the two of them, and the density of the notes' appearances on beat shows that dotted quavers always precede two*

demisemiquavers. The heat map shows that the pattern repeats itself in all the voices, suggesting that this is a fixed motif. We may presume that it is a part of the subject. We test the hypothesis by filtering by subject in the main dashboard's Piano Roll *chart (see 3.1).*

3.3 Comparative Analysis Dashboard

This dashboard is designed for the overall comparison of the musical form of different pieces in a corpus. It contains statistical information regarding ambitus, rhythm, number of notes and pauses, and rhythmic distribution. In the bottom half, there is a possibility to choose between three interactive charts: The first representation shows the number of intervals as a *column chart* where the mode of intervals is differentiated, the second contains summarized information about the relative frequency of intervals, and the last one is a rhythmic tree map.

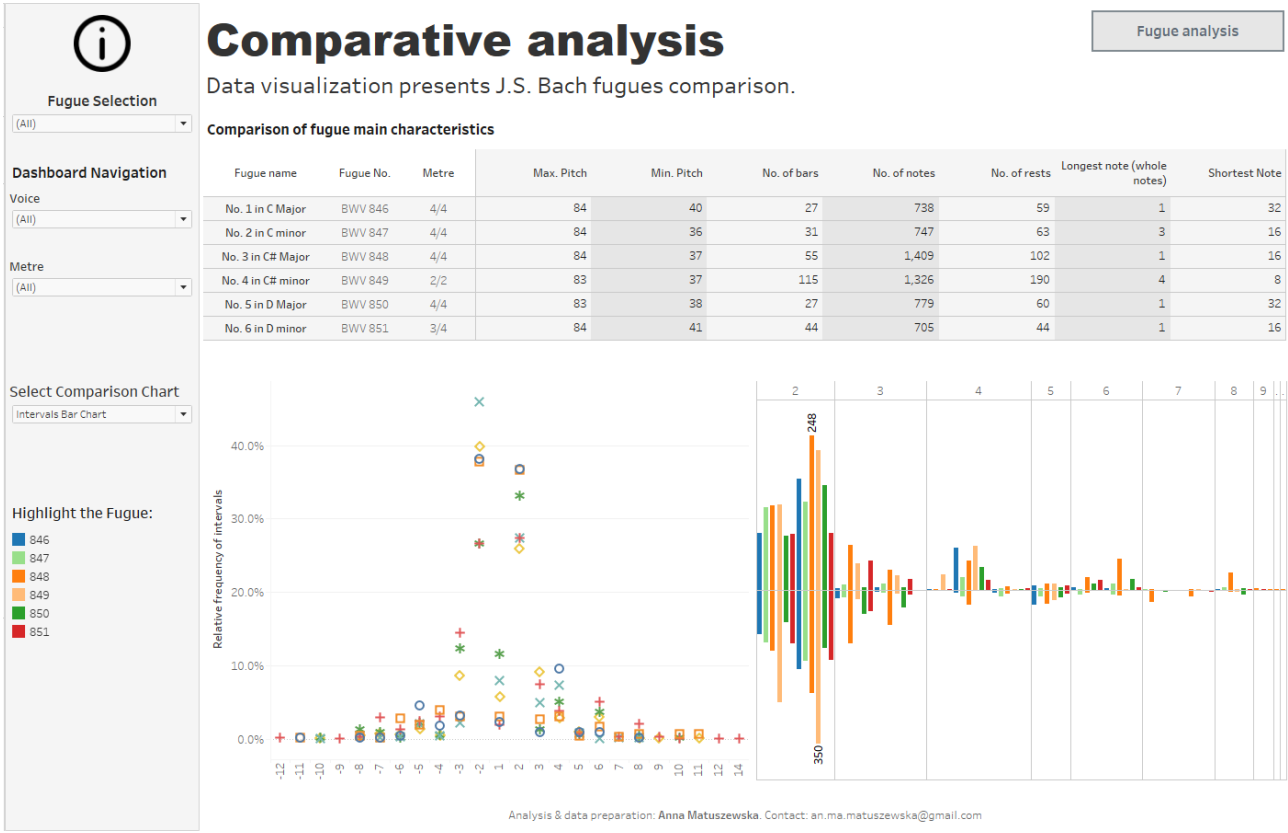


Figure 3: A view of the *Comparative Analysis* dashboard in the Fugue Analysis project showing the data visualization of main characteristics of J.S. Bach's fugues No. 1–6, BWV 846–851.

Conclusion

The application of diagrams in research creates cognitive advantages and enhances inferring. The main goal of this project was to adapt the diagrammatic thinking concept to music analysis and explore its potential to enrich research practice. Different representational formats of data presentation incorporated into each dashboard and their dynamic transformations creatively steer our minds to find a path to possible conclusion. The combination of visual and semantic data display makes diagrammatic representation efficient. Furthermore, the explanatory hypotheses generated for the chosen scope may be deductively proved, since exploring all proposed charts gives a detailed overview of chosen aspects; the change of the scope of analysis asserts conditions for inductive testing of the hypotheses.

The proposed approach of symbolic music data processing and visualization facilitates the comprehension of a single piece of music and music collections. Additional options give insight into the musical form from different angles and provide a space for exploration in accordance with abductive reasoning. Explanatory hypothesis inferred for one piece of music may be applied to other pieces from the corpora, and outcomes may be compared with the integrated music score.

In the light of the first experiences within this project, further developments towards tools for an automated process of data preparation and visualization are desirable to enable a fluid integration of diagrammatic reasoning and visually driven explorative processes into musicological research. The further development of visual analysis will be based on empirical observations of the users interacting with dashboards. Owing to this approach, our analysis method should be pragmatic and easy to implement in everyday practice.

Acknowledgments

This project is part of a Bi-nationally Supervised Doctoral project funded by the DAAD (German Academic Exchange Service).

Works Cited

[1] Bauer, Matthias, and Christoph Ernst. *Diagrammatik Einführung in ein kultur- und medienwissenschaftliches Forschungsfeld*. Bielefeld: transcript Verlag, 2010.

[2] Blackwell, Alan Frank. "Metaphor in Diagrams" University of Cambridge, Cambridge, 1998. PhD Dissertation, https://doi.org/10.17863/CAM.19818.

[3] Bruhn, Siglind. *J. S. Bachs Wohltemperiertes Klavier. Analyse und Gestaltung*. Waldkirch: Edition Gorz, 2013

[4] Giardino, Valeria. "3. Behind the Diagrams: Cognitive Issues and Open Problems" in *Thinking with Diagrams. The Semiotic Basis of Human Cognition*, ed. Sybille Krämer and Christina Ljungberg. Berlin, Boston: De Gruyter Mouton, 2016, 77–102, https://doi.org/10.1515/9781501503757-004..

[5] Hofmann, David M. "Music Processing Suite". https://www.musicprocessing.net/ (accessed January 12, 2022).

[6] Hofmann, David M. "Music Processing Suite: A Software System for Context-based Symbolic Music Representation, Visualization, Transformation, Analysis and Generation" University of Music, Karlsruhe, 2018. PhD Dissertation.

[7] Krämer, Sybille, and Christina Ljungberg. "Thinking and Diagrams – An Introduction" in *Thinking with Diagrams. The Semiotic Basis of Human Cognition,* ed. Sybille Krämer and Christina Ljungberg. Berlin, Boston: De Gruyter Mouton, 2016, 1–20, https://doi.org/10.1515/9781501503757-001.

[8] Nepomuceno-Fernández, Ángel, Fernando Soler-Toscano, and Fernando R. Velázquez-Quesada. "The Fundamental Problem of Contemporary Epistemology" *Teorema* 33, no. 4 (2014), 89–103.

[9] Peirce, Charles S. "Lectures on Pragmatism. Lecture VI: Three Types of Reasoning" in *The Collected Papers of Charles Sanders Peirce [1866–1913]*, vol. 5: *Pragmatism and Pragmaticism.* Electronic edition reproducing vols. I–VI ed. Charles Hartshorne and Paul Weiss (Cambridge: Harvard University Press, 1931–1935); vols. VII–VIII ed. Arthur W. Burks (same publisher, 1958). Charlottesville: Intelex Corporation, 1994, https://web.archive.org/web/20210508045931/https://www.textlog.de/7658.html.

[10] Sapp, Craig. "Kern Scores". 2001. http://kern.ccarh.org/ (accessed January 12, 2022).

[11] Sapp, Craig. "Verovio Humdrum Viewer". 2017. http://verovio.humdrum.org/ (accessed January 12, 2022).

[12] Stjernfelt, Frederik. *Diagrammatology. An Investigation on the Borderlines of Phenomenology, Ontology, and Semiotics* (Synthese Library 336). Dordrecht: Springer Netherlands, 2007, https://doi.org/10.1007/978-1-4020-5652-9.

[13] Urbański, Mariusz. *Rozumowania abdukcyjne. Modele i procedury*, Poznań: Wydawnictwo Naukowe UAM, 2009.

Diagrammatic analysis of J.S. Bach's The Well-Tempered Clavier fugues BWV 846-851

Anna Matuszewska | The Institute of Literary Research of The Polish Academy of Sciences | an.ma.matuszewska@gmail.com

Christoph Seibert | Institute for Music Informatics and Musicology University of Music Karlsruhe | christoph.seibert@hfm-karlsruhe.de

INTRODUCTION

Digital Tools facilitate searching through large music corpora and serve music analysis well. As a result, the analysis of large corpora has been simplified, but establishing relations between both subsets of a corpus and particular pieces of music can still be perplexing. In order to make computational analysis outcomes more explicit, comprehensible, and flexible, the visual exploration of analytical data should be regarded as a critical part of the research process.

Symbolic music data can be presented on different kinds of diagrams, thus becoming part of the epistemic process. Thinking with diagrams as hybrid forms of word and image forms the fundamental semiotic basis of human cognition [1]. Diagrams are 'scaffoldings' enhancing the ability to reason and infer [2].

THE AIM OF THE PROJECT

The main aim of our project is to examine possibilities of displaying analysis results to enable interaction with data and ensure conditions for abduction under assumptions of diagrammatic reasoning as proposed by Charles Sanders Peirce, not a thorough analysis itself.

METHODOLOGY

- Music information from fugues BWV 846-851 composed by J.S. Bach was retrieved with the Humdrum Tools [3] and Music Processing Suite (MPS) [4][5] software.
- Analytical information was processed and translated into a relational database.
- The visual display of the retrieved information was accomplished with dashboards using the data visualization software Tableau Public [6].
- The outcomes of the music analysis are presented in diverse representational formats on a group of combined interactive charts that are filtered jointly.

DIAGRAMMATICS

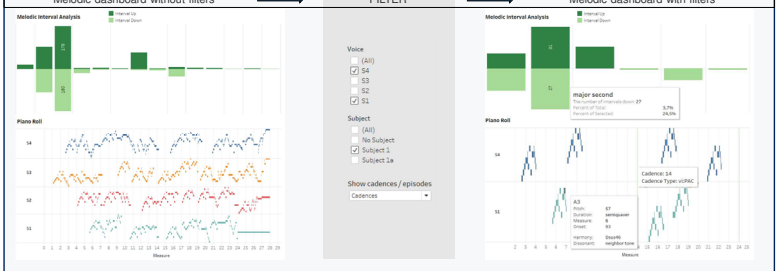
The concept of diagrammatics refers to the heritage of the American pragmatist and semiotic Charles S. Peirce. It addresses the design and processing of knowledge in diagrams. In Peirce's understanding it puts in the spotlight an epistemological process primarily associated with mentally realized diagrammatic conclusions. He stated that conclusive thinking is written diagrammatically and interlocked diagrammatic reasoning with abduction [7].

> *Abduction is the process of building an explanatory hypothesis. […] Deduction proves that something must be; Induction shows that something actually works; Abduction only suggests that something can be. Its only justification is that deduction can turn its assumption into a prediction which induction can test, and that if we can ever learn anything or understand phenomena at all, this can only be achieved through abduction. [8]*

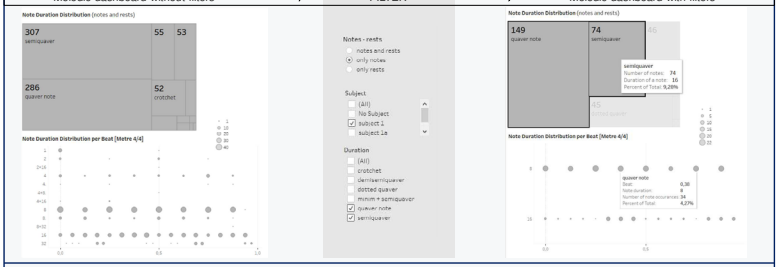
Interactive Elements for Visual Exploration

Detailed analysis of a fugue (example: Fugue BWV 846 in C Major)

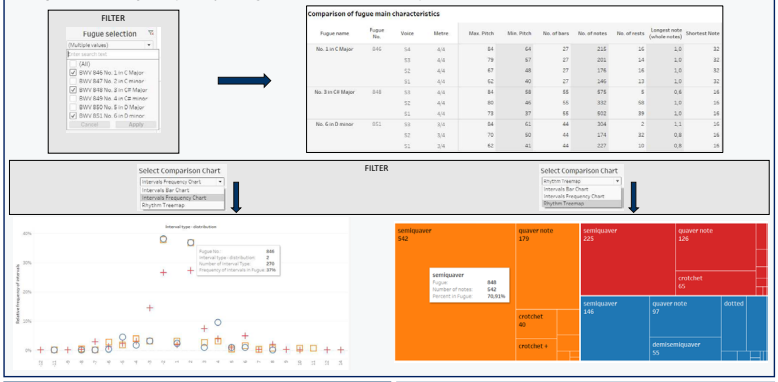
Melodic dashboard without filters → FILTER → Melodic dashboard with filters



Melodic dashboard without filters → FILTER → Melodic dashboard with filters



Comparative analysis (example: fugues BWV 846, 848, 851)



CONCLUSION

The application of diagrams in research creates cognitive advantages and enhances inferring. Different representational formats of data presentation incorporated into each dashboard, and their dynamic transformations creatively steer our minds to find a path to the conclusion. Furthermore, the explanatory hypothesis generated for the chosen scope may be deductively proved since exploring all proposed charts gives a detailed overview of chosen aspects, the change of the scope of analysis asserts conditions for inductive testing of the hypothesis.

The proposed approach of symbolic music data processing and visualization facilitates the comprehension of a single piece of music and music collections. Additional options give insight into the music form from different angles and provide a space for exploration in accordance with abductive reasoning. Explanatory hypothesis inferred for one piece of music may be applied to other pieces from the corpora, and outcomes may be compared with the integrated music score.

BIBLIOGRAPHY

[1] Krämer, Sybille and Christina Ljungberg. "Thinking and diagrams – An introduction" in *Thinking with Diagrams. The Semiotic Basis of Human Cognition*, ed. Sybille Krämer and Christina Ljungberg. De Gruyter Mouton, 2016, 1-19.

[2] Giardino, Valeria. "3. Behind the diagrams: cognitive issues and open problems" in *Thinking with Diagrams. The Semiotic Basis of Human Cognition*, ed. Sybille Krämer and Christina Ljungberg. De Gruyter Mouton, 2016, 77-102.

[3] "Humdrum Tools" (2021). https://www.humdrum.org/tool/

[4] Hoffmann, David. "Music Processing Suite" (University of Music Karlsruhe 2021). https://www.musicprocessing.net/

[5] Hofmann, David M. "Music Processing Suite: A Software System for Context-based Symbolic Music Representation, Visualization, Transformation, Analysis and Generation." University of Music, Karlsruhe, 2018. PhD Dissertation.

[6] "Tableau Public" (Tableau Software, 2021). https://public.tableau.com/en-us/s/

[7] Bauer, Matthias and Christoph Ernst. „Diagrammatik Einführung in ein kultur- und medienwissenschaftliches Forschungsfeld", transcript Verlag, 2010, p.64.

[8] Peirce, Charles Sanders. „Vorlesungen über Pragmatismus", ed. Elisabeth Walther, 1991, p.115.

FURTHER INFO

1. https://public.tableau.com/app/profile/anna.m8322/viz/poster-v3/1-MAIN
2. https://github.com/AnnaMariaMat/MEC-2021-Bach