Schüler, Nico. "Towards a History and Evaluation of Statistical and Information-Theoretical Analysis of Melodic Incipits," *Theoria: Historical Aspects of Music Theory* 13 (2006): 113-126.

Towards a History and Evaluation of Statistical and Information-Theoretical Analysis of Melodic Incipits

Nico Schüler
Texas State University-San Marcos

1. Introduction

Mathematical or computer-assisted music analysis provides analytical tools to help solve problems which cannot be solved (sufficiently) with traditional methods of music analysis. For instance, it may clarify stylistic characterizations and questions of unclear authorship, it helps investigate (historical) musical developments, it is useful for developing new theoretical systems, for research on acoustics and performance, as well as for cognitive and artificial intelligence research. Mathematical approaches to music analysis go back hundreds of years, but experienced a boom in the early 20th century. Computer-assisted music analysis, on the other hand, started in the mid-1950s (Schuler 2000, Schüler 1996).

Historically, many attempts to analyze melodies or parts of compositions mathematically or with a computer have concentrated only on the beginnings—incipits—of the music, making these incipits the object of their analyses as a representative of entire pieces. This was usually justified by the large number of calculations or by the insufficient memory capacity of the computer. Scholars believed that these incipits are sufficient for representing entire pieces of music. However, the authors usually did not reflect upon the possible effects that the length of the musical excerpts has on the analytical results; they did not question the validity of the analytical results. This paper will first summarize the use of incipits in the history of mathematical—including computer-assisted—music analysis. Then, using the methodology of falsification, analytical results will show the differences of statistical and information-theoretical results, when different incipit lengths are being used to analyze music. This case study is based on analyses of divertimentos for three bassett horns by Wolfgang Amadeus Mozart (1756-1791).

2. The Use of Incipits and Excerpts in the History of Mathematical Music Analysis

Many approaches of mathematical or computer-assisted music analysis, especially those applied to folk songs, followed Béla Bartók's methods of analysis and classification. Bartók's cataloging of East-European melodies was not only based on melody and text incipits, but also on different inner-musical characteristics. Bartók's methods of classification were then further developed by Alica Elschekova (1966, 1975) and others. Hereby, 'classification' means the grouping of music into specific categories, based on specific characteristics. The main focus of mathematical and computer-assisted analysis of folk music was the search for rules that put melodies into a specific category and the search for procedures that determined melodic variants. This, in turn, led to the development of analytical methods that investigated single or combined characteristics of music.

Similar attempts have been made to characterize art music. Characteristics for musical classification are, for instance, based on the following statistical and information-theoretical measurements:

> *Arithmetic Mean*: The arithmetic mean (sometimes called 'average') is calculated by dividing the sum of all elements (e.g., pitches, coded as a numerical value) by the number of elements.
>
> *Chi Square Test for Goodness of Fit*: The Chi Square Test can calculate whether two samples (e.g., melodies) are equal or not. Thus, it compares observed and expected frequencies.
>
> *Entropy*: Entropy is a form of measurement found in the conceptual methodology of information theory and is not related to semantics, but to syntax. It is an index of the degree of 'information' found by analyzing single elements (e.g., pitches or tone durations) or groups of elements taken as a unit. In the latter case, the entropy is of 'higher order'. The entropy is specifically the negative sum of all logarithms of the probability of each event multiplied by the probability of each event. (Shannon and Weaver 1949, p. 49f.) The average entropy of a melody, for instance, is the negative sum of all logarithms of the

probability of each note multiplied by the probability of each note. In case of calculating the entropy of the second order, the specific succession of two notes are seen as one element.

Frequency: There are two types of frequencies: absolute frequency and relative frequency. Absolute frequency is the exact number of a specific class of elements (e.g., pitch class c), while relative frequency is the absolute frequency of a specific element related to the total number of elements. The relative frequency is always smaller or equal to one, because the denominator is always larger than, or equal to, the numerator. The quotation in percent results from the multiplication with the factor 100.

Standard Deviation and Variance: Standard Deviation and Variance give information about the distribution of the elements (e.g. pitches, tone durations, or intervals) around the mean, i.e. the average distance of all elements from the mean. The Variance is calculated by permanently subtracting the mean from each element, squaring all results, adding them together and dividing them by the total number of all elements minus one. The Standard Deviation is the square root of the Variance.

Transition Frequency and Transition Probability: Transition Frequency is the frequency with which certain elements (e.g., pitches) occur in some places, when it is known that certain others occur in previous places. Transition Probability is the probability of an element (e.g., a note or a group of notes) which follows another specific element (note or group of notes).

Several examples will demonstrate the historical use of applying incipits or excerpts with the measurements mentioned above.

Based on communication theory, William J. Paisley (1964, 1969) made a fundamental contribution to identifying authorship in music (and with that, stylistic characteristics) by exploring "minor encoding habits", i.e. details in works of art (which would be, for instance, too insignificant for imitators to copy). To take an example from a different field, master paintings can be distinguished from imitations by examining details, such as the shapes of fingernails. Similarly, Paisley showed that there are indeed significant minor encoding habits in music. He analyzed note-to-note pitch transitions in the ***first six***

notes of each of the 320 themes by Johann Sebastian Bach, Joseph Haydn, Wolfgang Amadeus Mozart, Ludwig van Beethoven and Johannes Brahms. He chose the parameter 'pitch', because pitches can be easily coded for computer processing and because some research on tonal transitions had already been reported. In his first analysis, performed on the Stanford University 7090 computer, Paisley (1969) calculated interval frequencies of up to six semitones within the first 6 notes of two 160-theme-samples. Furthermore, he calculated the Chi Square Test for Goodness of Fit of those interval distributions for the two samples. While these results could not significantly distinguish Haydn, Mozart, and Beethoven, Paisley claimed a successful distinction between these composers with his second analysis, in which he calculated frequencies and chi squares of two-note transitions between the classes tonic, third, fifth, all other diatonic tones and all chromatic tones. In both analyses, the results of the chi square test were then compared with results from "unknown samples" (Mozart, Beethoven, Georg Friedrich Händel, and Felix Mendelssohn). The results from analyzing incipits of themes by Mozart and Beethoven could (in the second analysis) be successfully matched with the "known" Mozart- and Beethoven-samples, while Händel and Mendelssohn were significantly different. But even though only a modest amount of data was involved in this investigation, and even though a reduction of the number of possible intervals to seven (based on inversions as well as on neglecting the direction) seems to be questionable, Paisley's study was well documented and its results were, considering the time of the study, very impressive. Several other authors referred later to Paisley's approach.

In his studio for experimental music at the University of Illinois, Lejaren A. Hiller collaborated in several analytical research projects. One of the projects—dissertation research conducted by Calvert Bean (Bean 1961, Hiller 1964, Hiller 1966)—involved a comparison of four sonata ***expositions*** (by Wolfgang Amadeus Mozart, Ludwig van Beethoven, Paul Hindemith, and Alban Berg), mainly based on first-order entropies of pitches and intervals as well as on the "speed of information" (i.e., of entropy), which was calculated via note density and tempo.

During the 1970s, Lynn Mason Trowbridge used incipits to analyze Burgundian Chansons (Trowbridge 1970) and polyphonic compositions of the Renaissance (Trowbridge 1971).

Dean Keith Simonton's research was also based on Paisley's analytical attempts (Simonton 1980). Simonton combined computer-assisted analyses of two-note transitions within the ***first 6 notes*** of 5046 classical themes (by ten well-known composers) with broader, more encompassing, analyses of psychological and socio-cultural factors. His goal was to find musical characteristics that make a musical theme 'famous'. 'Thematic fame' was defined, on the one hand, with regard to the frequency of performances, recordings, and citations (Simonton 1980, p. 210). On the other hand, "melodic originality was operationalized as the sum of the rarity scores for each of the theme's 5 transitions" (Simonton 1980, p. 211). Chromaticism and dissonant intervals played an important role in the statistical calculations. But Simonton neither calculated note transitions of higher orders (beyond two-note transitions), nor did he calculate transitions related to duration or rhythm. Simonton's main results can be summarized as follows: 1. 'thematic fame' is a positive linear function of melodic originality; 2. melodic originality of themes increases over historical time; 3. melodic originality of a theme increases when composed under stressful circumstances in a composer's life; and 4. melodic originality is a curvilinear inverted backwards-J function of the composer's age. (Simonton 1980, pp. 213-215.) Even though some of his results are still valid, most of them are not, especially those dealing with the empirical determination of 'thematic fame' and with the correlation of 'creativity' and Simonton's calculations of 'melodic originality' (interpreted as 'novelty'). Recent research on musical creativity (e.g., Deliège and Wiggins 2006) does not support Simonton's understanding of 'melodic originality'. Although, within a history of computer-assisted music analysis, the attempt of combining psychological and socio-cultural factors and statistical analyses was an important step, the use of short incipits was most questionable.

Another example from the 1980s, using incipits for stylistic characterization of art music based on statistical analysis, is the research by Alison Crerar (1985). Crerar analyzed 105 ***incipits*** of compositions by Valentini, Arcangelo Corelli, Antonio Vivaldi, Johann Sebastian Bach and Ludwig van Beethoven (especially with the goal to compare Valentini with Corelli and Vivaldi), following W. J. Paisley's earlier research. After refining and extending Paisley's procedures by statistical calculations of pitch, intervals, and scale degrees etc., Crerar showed that it is possible thereby to distinguish between the works of different composers and to clarify the authorship of specific compositions.

3. Evaluation of the Method of Using Incipits and Melodies of Various Lengths

The goal of the following study was to evaluate the method of using incipits as representatives of entire compositions—a method that has occupied a prominent position in the history of mathematical and computer-assisted music analysis. The main methodological approach taken here (in evaluating the analytical procedure) is falsification. Falsification, the act of showing one instance of something to be false or erroneous to reflect on the potential of a theory, is a powerful tool for evaluating methods of mathematical and / or computer-assisted music analysis for two reasons: it neither requires analyzing a large number of compositions nor carrying out extensive verification.

This study is based on analyses of Wolfgang Amadeus Mozart's *25 Pieces (five Divertimenti) for Bassett Horns* KV 439b (Mozart 1975). Limiting the pieces to be analyzed to those of a specific composer and in a specific genre is necessary to eliminate distinguishing musical characteristics that are deduced by the following stylistic differences:

- the differences between styles from different periods,
- the differences between different personal styles, and
- the differences between different genres.

Focussing on analyzing only *one* set of compositions (in the same genre) allows one to reduce the probability of error, which could occur when different characteristics of style or genre influence the outcome of statistical and information-theoretical analyses. Analyses that focus on differences between genres, personal styles, or time periods can only be carried out *after* successfully applying certain measurements to analyzing music with a reduced number of distinguishing characteristics. If such an analysis with a reduced number of distinguishing characteristics did not precede, characteristics of time, style or genre can hardly be distinguished, i.e. personal style, for instance, can influence analyses of genre characteristics, and so forth (Schüler 1992).

Mozart's *25 Pieces (five Divertimenti) for Bassett Horns* KV 439b were composed in 1783; the original instrumentation is not certain. With the selection of divertimenti, a musical form was chosen that was historically a continuation of the suite; the character of the divertimento belongs to *Gebrauchsmusik*. All five divertimenti in this group of compositions have five movements each. For the purpose of this study, the *Neue Mozart Ausgabe*, Serie VIII (Kammermusik), Werkgruppe 21 (Duos und Trios für Streicher und Bläser) was used for the analyses (Mozart 1975).

The computer program used in this study is *MUSANA*—a program developed by Nico Schüler and the German physicist Dirk Uhrlandt (Schüler and Uhrlandt 1994/1996). *MUSANA*, written in the programming language Turbo-

Pascal, is a music analysis program that draws on statistics and information theory (Uhrlandt and Schüler 1992). *MUSANA* extends traditional methods of music analysis by computer-assisted methods, it does not replace them. It is crucial for the outcome of computer-assisted analysis to integrate both, computer-assisted and traditional methods of music analysis.

This study compares analyses performed with different lengths of the excerpts (incipits). Each part of the first movements (Allegro) of both, Divertimenti I and II, of KV 439b is analyzed in the following lengths:

- only the first 10 notes (and rests)
- only the first 20 notes (and rests)
- only the first 40 notes (and rests)
- only the first 60 notes (and rests)
- the entire piece.

The task is to compare the following statistical values, most often used in the past to supposedly characterize a certain musical style:

- average pitch (using the internal numerical code [of MUSANA] and considering the duration of each note) and its standard deviation
- the average interval size (disregarding the direction; half step = 1)
- the average tone duration (statistically, here, as a partial of a whole note)
- first order entropy (considering the duration of each note, not just their number of appearances)

The MUSANA results of the analyses are as follows (see the following pages):

Table 1: Allegro from Divertimento I, upper voice

	10 notes	20 notes	40 notes	60 notes	All (834)
Average Pitch & Standard Deviation	57.0 ± 2.7 (g#')	58.2 ± 3.6 (a')	57.8 ± 3.0 (a')	58.4 ± 3.4 (a')	58.2 ± 3.6 (a')
Average Interval Size	0.8	1.1	1.8	2.1	2.4
Average Tone Duration & Standard Deviation	0.2125 ± 0.0976	0.1776 ± 0.0843	0.2279 ± 0.2089	0.2255 ± 0.1836	0.2386 ± 0.2106
First Order Entropy	0.95604	1.65300	1.89431	2.01050	2.47483

Table 2: Allegro from Divertimento I, middle voice

	10 notes	20 notes	40 notes	60 notes	All
Average Pitch & Standard Deviation	48.5 ± 3.7 (c#')	52.1 ± 5.8 (d#')	53.4 ± 4.1 (e')	53.6 ± 4.2 (f')	52.9 ± 4.2 (e')
Average Interval Size	1.3	1.4	1.7	1.9	2.5
Average Tone Duration & Standard Deviation	0.2125 ± 0.0976	0.1776 ± 0.0843	0.2286 ± 0.1650	0.2052 ± 0.1418	0.1922 ± 0.1302
First Order Entropy	0.95604	1.75553	2.00176	2.06143	2.47987

Table 3: Allegro from Divertimento I, lower voice

	10 notes	20 notes	40 notes	60 notes	All
Average Pitch & Standard Deviation	36.9 ± 5.3 (c)	36.7 ± 5.8 (c)	41.1 ± 6.6 (e)	38.7 ± 7.0 (d)	41.6 ± 6.1 (f)
Average Interval Size	4.0	4.5	3.0	5.2	3.8
Average Tone Duration & Standard Deviation	0.2125 ± 0.0976	0.1964 ± 0.0911	0.2311 ± 0.0821	0.2170 ± 0.1333	0.1976 ± 0.1258
First Order Entropy	1.21820	1.46880	2.35662	2.20022	2.71846

Table 4: Allegro from Divertimento II, upper voice

	10 notes	20 notes	40 notes	60 notes	All (416)
Average Pitch & Standard Deviation	57.1 ± 3.7 (g#')	57.6 ± 3.5 (a')	57.8 ± 3.3 (a')	59.2 ± 3.8 (a#')	58.8 ± 3.8 (a#')
Average Interval Size	2.8	2.2	2.1	2.1	2.4
Average Tone Duration & Standard Deviation	0.2222 ± 0.1534	0.1776 ± 0.1169	0.1757 ± 0.1179	0.178 ± 0.1332	0.1812 ± 0.1368
First Order Entropy	1.95212	2.04701	2.09673	2.35861	2.40463

Table 5: Allegro from Divertimento II, middle voice

	10 notes	20 notes	40 notes	60 notes	All (410)
Average Pitch & Standard Deviation	48.1 ± 2.2 (b)	51.0 ± 4.3 (d')	50.0 ± 4.0 (c#')	51.8 ± 4.3 (d#')	52.3 ± 4.6 (d#')
Average Interval Size	1.9	1.8	2.0	2.0	2.3
Average Tone Duration & Standard Deviation	0.2344 ± 0.1159	0.1806 ± 0.0952	0.1786 ± 0.0958	0.1838 ± 0.1190	0.1826 ± 0.1231
First Order Entropy	1.68077	2.27938	2.34430	2.3097	2.51913

Table 6: Allegro from Divertimento II, lower voice

	10 notes	20 notes	40 notes	60 notes	All (486)
Average Pitch & Standard Deviation	40.8 ± 3.6 (e)	41.4 ± 3.4 (e)	41.2 ± 3.2 (e)	40.1 ± 3.6 (d#)	39.9 ± 4.9 (d#)
Average Interval Size	1.3	1.8	1.9	2.3	3.0
Average Tone Duration & Standard Deviation	0.1250 ± 0.0000	0.1324 ± 0.0294	0.1326 ± 0.0298	0.1297 ± 0.0238	0.1288 ± 0.0240
First Order Entropy	0.32508	1.18372	1.13290	1.53338	2.14939

Evaluative results of the few calculations presented above provide an astonishing picture of the value of such calculations, at least taken separately, i.e. not as one component of more complex statistical measurements, such as multi-variate analysis, cluster analysis, or factor analysis. The interpretation of the test results can be summarized as follows:

- While the mean values of pitch do not vary much, their standard deviations may vary by more than 100%. In the middle voice of the Allegro from Divertimento II, for instance, the standard deviation from the pitch average for the first 10 notes is 2.2, but for the entire piece

- Similarly, the average interval size varies considerably. The values for shorter incipits (10 and 20 notes / rests), in particular, are far from being close to the average of the entire voice. The upper voice of the Allegro from Divertimento I, for instance, shows 0.8 as the average interval size for the first 10 notes and 1.1 for the first 20 notes, but the average interval size of the entire voice is 2.4, i.e. three times more than the value for the first 10 notes.

- Not only can incipits not accurately characterize the entire piece, but even the values of the same parts (voices) within different pieces are not comparable. The average tone durations of the lower voices of both Allegros are 0.1976 and 0.1288, respectively—a difference of more than a sixteenth note.

- The calculations with regards to the lower voice of the Allegro from Divertimento II demonstrate the falseness of the assumption that an incipit's standard deviation from the average tone duration can be used to characterize a larger part of the piece or the whole piece. While the standard deviation of the first 10 notes is zero, the standard deviation of the first 20 notes is already 0.0294.

- First-order entropies seem not to be significant for a 10-note incipit. The entropies of all larger excerpts show a natural, and consistent, growth when the incipits become longer.

Conclusions

Using incipits as if they were representative of the whole composition has been common practice in mathematical and computer-assisted music analysis since its beginning. Since evaluations of this practice have not been available, even during the 1990s, calculations based on incipits were put forth as adequate characterizations of compositions or even of a composer's style. The data provided in this study clearly show that there is no statistical basis for the assumption that incipits have a sufficient size for discriminatory tasks or style characterizations.

This study shows that scholars have not been critical enough in the past with the method of analysis. Many other methods of (mathematical and computer-assisted) music analysis are in urgent need of evaluation. A methodological approach that uses falsification in the manner demonstrated in this study provides a powerful way to evaluate methods of mathematical and computer-assisted music analysis. However, many questions remain about verification or falsification of analytical methods:

- To which kind of music are the chosen methods of analysis applicable?
- Using a specific method of analysis, which musical characteristics can influence the analytical results?
- How does each musical characteristic influence the analytical results?
- How can we separate those musical characteristics that influence the analytical results? -Which of these musical characteristics are most influential for the chosen method of analysis?
- Is it possible to weight the musical characteristics in order to receive more objective analytical results?
- Which methods of music analysis are less useful and can be eliminated?
- How can we design a more interactive process of analysis, so that traditional methods of music analysis and mathematical or computer-assisted methods of music analysis can merge in more useful ways?

Bearing in mind that all analytical results are influenced by the method used, answering all of these and similar other questions can help improve methods of mathematical and computer-assisted music analysis (Schüler 2002).

References

Bean, Calvert. Information Theory Applied to the Analysis of a Particular formal Process in Tonal Music. Ph.D. dissertation. University of Illinois, 1961.

Crerar, Alison. "Elements of a Statistical Approach to the Question of Autorship in Music," Computers and the Humanities XIX/3 (1985): 175-182.

Elscheková, Alica. "Methods of Classification of Folk Tunes," Journal of International Folk Music Council XVIII (1966): 56-76.

Elscheková, Alica. "Systematisierung, Klassifikation und Katalogisierung von Volksliedweisen," Handbuch des Volksliedes, ed. by R. W. Brednich, L. Röhrich, and W. Suppan, vol. 2. München: Fink, 1975. pp. 549-582.

Deliège, Irène, and Geraint A. Wiggins. Musical creativity: Multidisciplinary Research in Theory and Practice. New York: Psychology Press, 2006.

Hiller, Lejaren A. Informationstheorie und Computermusik. Darmstädter Beiträge zur Neuen Musik 8. Mainz: B. Schott's Söhne, 1964.

Hiller, Lejaren A., and Calvert Bean. "Information Theory Analyses of Four Sonata Expositions," Journal of Music Theory X (1966): 96-137.

Mozart, Wolfgang Amadeus. "Fünfundzwanzig Stücke (fünf Divertimenti) für drei Bassetthörner (KV 439b)," Neue Ausgabe Sämtlicher Werke. Serie VIII: Kammermusk. Werkgruppe 21: Duos für Streicher und Bläser, ed. by Dietrich Berke and Marius Flothuis. Kassel: Bärenreiter, 1975. 67-119.

Paisley, William J. "Identifying the Unknown Communicator in Painting, Literature and Music: the Significance of Minor Encoding Habits," Journal of Communication XIV/4 (December 1964): 219-237.

———. "Studying 'Style' as Deviation from Encoding Norms," The Analysis of Communication Content. Developments in Scientific Theories and Computer Techniques, ed. by George Gerbner, Ole R. Holsti, Klaus Krippendorff, William J. Paisley and Philip J. Stone. New York: Wiley & Sons, 1969. 133-146.

Schuler, Nico. Methods of Computer-Assisted Music Analysis. History, Classification, and Evaluation. Ph.D. dissertation. East Lansing: Michigan State University, 2000.

———. "Methoden und Resultate einer computergestützten Musikanalyse Mozartscher Werke," Wolfgang Amadeus Mozart. Referate des wissenschaftlichen Kolloquiums der Greifswalder Mozart-Tage am 3. Dezember 1991, ed. by Nico Schüler and Lutz Winkler. Greifswald: Institut für Musikwissenschaft, 1992. 103-120.

———. "Methoden computerunterstützter Musikanalyse - ein historischer Überblick," Zum Problem und zu Methoden von Musikanalyse, ed. by Nico Schüler. Hamburg: von Bockel, 1996. 51-76.

———. Ed. Computer-Applications in Music Research: Concepts, Methods, Results. Frankfurt, New York: Peter Lang, 2002.

Schüler, Nico, and Dirk Uhrlandt. MUSANA 1.0 / 1.1 - ein Musikanalyseprogramm. Programmdokumentation. Peenemünde: Dietrich, 1994 / 1996.

Shannon, Claude E., and W. Weaver. The Mathematical Theory of Communication. Urbana: University of Illinois Press, 1949.

Simonton, Dean Keith. "Thematic Fame and Melodic Originality: A Multivariate Computer-Content Analysis," Journal of Personality 48 (1980): 206-219.

Trowbridge, Lynn Mason. "The Burgundian Chanson: An Index and Analysis of Incipits," Computers and the Humanities IV/5 (1970): 343-344.

Trowbridge, Lynn Mason. "An Analytical and Indexing System for Incipits of Renaissance Polyphonic Compositions," Computers and the Humanities V/5 (1971): 310.

Uhrlandt, Dirk, and Nico Schüler. 1992. "Informationsmaße in der computergestützten Musikanalyse," Kammermusik Heute 2 (1992): 102-112.