Changing the musical object: Approaches to performance analysis

Nicholas Cook
AHRC Research Centre for the History and Analysis of Recorded Music (CHARM)
Royal Holloway, University of London

It is a simple statistical fact that, for most people across the world, music means performance, whether live or recorded, and not scores. Given that, and the fact that records are the primary documents of music as performance, it might be thought of as strange that there is such an imbalance in musicological research training between score- and recording-based source criticism. Indeed to many musicologists it might not be immediately obvious what recording-based source criticism might mean. So let me offer a concrete example.

In 1987 Pearl Opal produced a compact disc called *Alessandro Moreschi: The last castrato* (OPAL CD 9823). On this compact disc is a recording of Moreschi singing the "Crucifixus" from Rossini's *Petite messe solennelle*, made at the Vatican on 3 April 1902. Fred and Will Gaisberg had been sent round the world by the Gramophone Company to "acquire a catalogue of native recordings", and came to Rome hoping to record Pope Leo XIII. They were denied a papal audience, so instead they recorded Moreschi (who at that time directed the Sistine chapel choir), and the resulting recording was issued on the Red G & T's label. It is this which has been remastered on the Pearl Opal compact disc.

That much you can read from the compact disc insert. The interpretational problems start when you listen to the compact disc itself, which reveals a highly unfamiliar style of vocal production, characterized by ornamentation that might easily be mistaken for lack of control. But then, maybe what we are hearing *is* lack of control. This session was Moreschi's first experience of being recorded and the circumstances must have been unnerving. The piano accompanist was probably placed at the opposite end of the room, disrupting the performers' sense of ensemble. Moreschi will have sung into a large horn, probably being told to step into it for quiet or low passages and away for loud or high ones. Such disruption, coupled to the sense of the uncanny that marked many first encounters with recording technology, may have completely undermined Moreschi's composure. The performance that was recorded may have been wholly unrepresentative of how Moreschi normally sang.

Then again, the sound we hear may be far from the sound that was recorded on 3 April 1902. The process of transfer involves making a number of what can only be called interpretive decisions—what stylus to use, how much noise reduction to apply—but the most basic decision is how fast to play the record. (Recording and playing speeds at the turn of the 20th century were as variable as orchestral pitch in the 19th.) The compact disc liner, which refers to another recording of the "Crucifixus" which Moreschi made two years later, makes the problem obvious:

> The pitching of Moreschi's records presented us with some problems since no-one had the slightest idea what his voice ought to sound like. However we discovered that when we pitched both the 1902 and 1904 recordings of Rossini's Crucifixus at the score key of A♭, all the records made at both sessions fell into score keys or reasonable keys, and sound vocally correct to several musicians for whom they were played.

The problem, of course, is how the "several musicians for whom they were played" could know what was "vocally correct" in a castrato, since the production (if that is the right word) of castrati was made illegal in 1870. The only reliable way in which we could possibly know how music sounded a century ago, after all, is by means of recordings.

The circularity inherent in interpreting early recordings is not complete, of course. We have contemporary accounts of how castrati sounded (though words are rarely unproblematic guides to performance practice). We have the other recordings Moreschi made during the 3 April 1902 session, and matrix numbers (numbers stamped into the original discs) make it possible to reconstruct the sequence in which he made them and so judge whether Moreschi started nervously but gained confidence as the session continued. We can also compare the 1902 recording of the "Crucifixus" with the 1904 one; indeed we might conclude from the fact that it was rerecorded that the 1902 recording was considered substandard. (Then again, the "Crucifixus" was the first recording made in the 11 April 1904 session, so perhaps they started with what was seen as a particularly successful performance from 1902 in order to settle Moreschi down.) And these are, of course, all source-critical arguments. The point I am trying to make is not, then, that recordings can tell us nothing about historical performance style: it is that recordings are historical documents just as scores are, and as much in need of historically and technically grounded interpretation. In short, discographical source-critical skills should be seen as an essential part of musicological research skills training today.

If musicologists have been slow to exploit the potential of sound recordings as documents for the writing of music history, they can hardly be blamed. Working with recordings is much like working with scores would be if it were not for RISM and other initiatives which have built the basic infrastructure for most musicological research. Although some online discographies are just beginning to appear, and the catalogues of a few sound archives are available on the web, serious work on primary sources entails obtaining sometimes inaccessible hard-copy publications (for instance, pre-war record company catalogues or composer discographies), and once you have established that a given recording exists, or existed, you have the problem of tracking down a copy. Even then, proper interpretation may not be possible without access to the original recording documentation which, if it exists at all, is probably held in the record company's archives (if the record company still exists, or if you can establish which of the present-day majors it was incorporated into). There are, quite frankly, easier ways of doing musicology.

Yet I do not think it is primarily these practical difficulties that have held back the use of sound recordings as primary sources for the writing of music history. The more entrenched problems have been conceptual. Though this is not the place to present the argument in detail, musicology can be seen as a by-product of 19th-century European nationalism. That is not simply to say that it was motivated by the grand project of reconstructing, or inventing, national origins and so underpinning national identity: It is to say that musicology was largely modelled on the values and practices of the then dominant discipline in the human sciences, philology. Hence the traditional musicological emphasis on editing as a process of removing the encrustations of later interpretation in order to arrive back at the original; hence, too, the emphasis on musical works rather than the patterns of social usage through which they acquire meaning.

And most relevantly in the present context, hence the idea of music as in essence a form of writing. To see musical works as texts which can be "reproduced" in performance (the standard translation for Schoenberg's or Adorno's writings on performance) is in effect to see music as a branch of literature. A comparison with poetry makes the point: Poetry can be read aloud, but we do not normally think of it as a performing art, because we see its meaning as already inherent in the written text—reading it aloud is an optional extra. In the same way, musicology's scriptist orientation inhibits its ability to conceptualize music as the performing art we all know it is; and that, of course, includes musicologists. My claim is not that musicologists are not interested in performance, but that the conceptual apparatus built into the discipline during the century before last makes it hard for us to translate that interest effectively into our writing.

It may seem that I have been offering a very old-fashioned characterization of musicology, a discipline which was, after all, radically transformed in the 1990s, with issues of social meaning, gender, ideology and so forth being brought firmly onto the agenda, alongside an acceptance of the importance of popular as well as world music traditions. Yet the "New" musicologists' opening up of the discipline did less than might have been expected to place performance—and, as the primary texts of performance, recordings—at the heart of the discipline. One reason, perhaps, was the influence of Adorno: Social meaning was seen as encoded in the musical text, which turned interpretation into a process of decoding. Because early "New" musicologists, in particular, looked for social meaning in the text, they were not inclined to observe it in the interpersonal and social transactions which those texts prompted—that is, in the meanings that are constructed through the act of musical performance.

If we think of music in these latter terms, the result is an approach maximally opposed to the idea of performance as the reproduction of a text. Seen this way, a Mozart string quartet, for example, choreographs an ongoing series of social interactions: That is what it means to keep in time and play in tune. The musicians do not keep in time because each is playing to some objective, standardized beat, as when session musicians play to a click track, wearing headphones. The "time" that a string quartet keeps is something negotiated between the players in the course of performance (as well as the rehearsals that have likely preceded it); the objective time of twice as long and half as long that you see in the score is only an initial starting point, a framework within which negotiation takes place. In other words, musical time as performed and experienced is a social construction, and the act of making music together is an act of forging and maintaining community. There is a sense, then, in which the sound of a string quartet is community made audible, and my claim is that this is inseparable from what the music means to us.

In his essay "The grain of the voice", which compares the singing of Dietrich Fischer-Dieskau and Charles Panzera, Roland Barthes complained of language's incapacity to handle music. And he continued, "rather than trying to change directly the language on music, it would be better to change the musical object itself, as it presents itself to discourse":[1] That is an apt description of the kind of disciplinary refocusing that is entailed in locating the production of musical meaning in performance and not, or not simply, the score. My aim in this paper, then, is to set out a personal selection of possible approaches to music as performance rather than as writing, maintaining some balance between what might be termed technically and culturally oriented approaches, and drawing not only on some work I have done, but also on some work I have not. I begin with the use of empirical measurements, generally presented in graphic form, not because I wish to give priority to these methods, but because they represent the most visible aspect of much recent writing on performance—and also because they raise some basic problems that I see as inherent in such work.

Figures 1–3 illustrate two principal approaches, both based on the measurement of time. The first two are from an article by José Bowen and illustrate the way in which musicologically interesting conclusions can be drawn from very simple empirical data. Both relate to the first movement exposition of Beethoven's fifth symphony.[2] Figure 1 charts the duration of the exposition in almost 70 different recordings (in seconds) against the date when the recording was made: the trend line indicates that, on average, the exposition is getting shorter as average tempos increase. (One might ask how robust this trend is, given the degree of scatter in the data.)
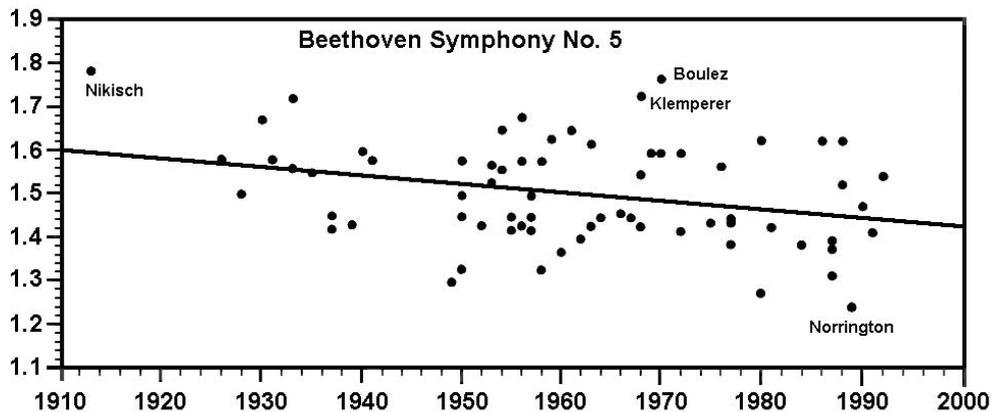

Fig. 1. José Bowen's analysis of durations in recordings of the first movement exposition from Beethoven's symphony no. 5.

Figure 2 also shows the duration of the exposition on the vertical axis, but now charts this against the average tempo of the first theme: The result is a measure of performance flexibility, understood as the extent to which the exposition as a whole takes longer ("relaxation") or shorter ("compression") than a metronomic extrapolation would suggest. The "Average Flexibility Line" shows how almost all performances are "relaxed",

[1] Roland Barthes, *Image, music, text*, ed. and trans. by Stephen Heath (London: Fontana, 1977) 180.
[2] José Antonio Bowen, "Tempo, duration, and flexibility: Techniques in the analysis of performance", *The journal of musicological research* 16/2 (1996) 115, 136.
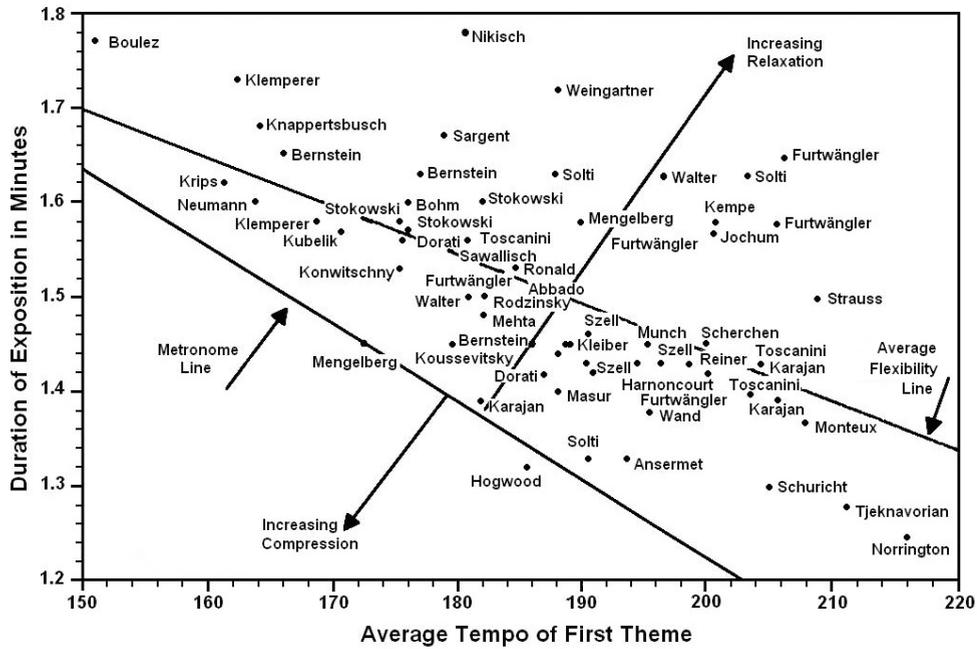
Fig. 2. José Bowen's flexibility graph of recordings of
Beethoven's symphony no. 5, first movement.

whether through taking the second theme more slowly or introducing rallentandos or rhetorical pauses; Hogwood's recording emerges as clearly anomalous.

The limitation of this approach is that so little musical information is embodied in it: duration of exposition and recording date in figure 1, duration of exposition and average tempo of first theme in figure 2. This means that, for example, the data conflate recordings which take the second theme slower, and recordings which introduce rallentandos or rhetorical pauses—which may be, musically, a very important distinction.
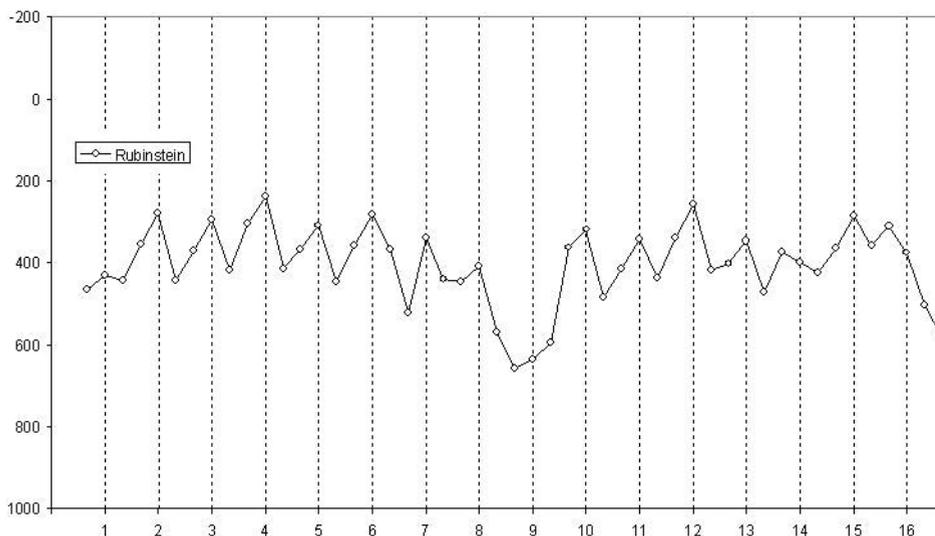
Fig. 3. Tempo graph of Artur Rubinstein's 1939 recording of
Chopin's mazurka op. 67, no. 4, bars 1–16 (milliseconds).

Hence the increasingly frequent use by musicologists of tempo graphs such as figure 3, a representation of bars 1–16 of Rubinstein's 1939 recording of Chopin's mazurka op. 67, no. 4 (a score is at figure 4). Such graphs are commonly generated by tapping on a computer keyboard in time to the music and importing the data into a spreadsheet program; in this case the data were made more accurate through subsequent visual comparison with the waveform. Since higher means faster and lower means slower (the vertical scale is milliseconds), the most obvious—and unsurprising—feature is the broad rallentando which marks the end of the first eight-bar phrase; Rubinstein's prolongation of the third beat of bar 6 audibly initiates the cadential motion. The almost identical profile of bars 2–4 and 10–12 reflects their musical identity (though, of course, Rubinstein did not have to play them the same way), with the downbeats of bars 4 and 12 being the shortest within these 16 bars. And the succession of V-shaped contours shows how consistently Rubinstein makes the first beat of the bar the shortest and the second the longest, with the third beat coming somewhere in the middle. This is, after all, a mazurka.

Fig. 4. Chopin's mazurka op. 67, no. 4, bars 1–16.

When musicologists and particularly analysts use graphs of this kind, they are usually trying to make some point about the interpretation of higher-level structure, as seen in terms of Schenkerian analysis, for instance. That's where the first of the basic problems I referred to comes in. If you begin, as people usually do, by analysing the score, and then see how far you can map the score-based analysis onto performance features, you are in effect filtering the performance data, discarding data that do not fit—or, at least, do not bear upon—the score-based analysis. If this nevertheless seems a plausible thing to do, that reflects how deeply embedded the idea of the peformance "reproducing" the compositional text remains: We may not use that word nowadays, but the idea persists

behind common terminologies of performance "projecting", "articulating", or "expressing" structure (the last term is particularly revealing, since it coincides with the Schenkerian model according to which surface design "expresses" structure). Working from page to stage, as they say in theater studies, such analyses do not so much change so much the musical object as pursue the old object by other means.

This is linked to the second problem, which is that—in contradistinction to the Bowen-style approach—such analyses often involve working with very few recordings of a given work, perhaps only one. Again it is the paradigm of reproduction that makes this seem plausible: The recording is interpreted directly in relation to the score, almost as if the music was being performed for the very first time. But in reality—and this has become even more the case since historical recordings have become widely disseminated— performers perform and listeners listen relationally, that is, in terms of patterns of similiarity and difference as compared with other performances. The "vertical" dimension which relates score to performance is, in this way, complemented by a "horizontal" dimension that relates each performance to others, and this second dimension can only be incorporated into the analysis if we analyse recordings comparatively, and use large enough data sets to be able to extrapolate trends from them.

And that leads to the third problem, which is the established musicological and analytical focus on the work as the essential unit of musical meaning. As in the case of the vertical and horizontal dimensions to which I have just referred, this is a matter of both/ and, not either/or: It makes sense to understand a given performance of op. 67, no. 4, in the context of other performances of op. 67, no. 4, but it also makes sense to understand it in terms of features—melodic, motivic, harmonic, rhythmic, combinations of all these things—that are shared between op. 67, no. 4, and other early mazurkas (op. 67, no. 4 is a posthumous work), the mazurkas in general, Chopin's piano music as a whole, and so on. Though individual works may establish their own performance traditions and, to that extent, have independent histories, it is absurd to imagine a stylistic history of performances of op. 67, no. 4, that is unrelated to performance style at these larger levels, and, accordingly, the analytical basis for such a history should involve features that cut across individual works. But that is where traditional analytical approaches are not well adapted for the purpose at hand. I can draw an analogy from computing: Microsoft Office makes it impossible to carry out complex operations across files, and hence the file becomes a basic conceptual framework in the same way that musical works are for most established analytical methods. Unix, by comparison, leads you to think in terms of patterns of data irrespective of their location within files: In this respect, it is much more similar to the style-analytical approaches which lay at the heart of musicology during the first half of the 20th century, but fell out of favor after World War II—ironically, not long before developments in computer technology made analysis based on large bodies of data fully practical for the first time. I see the rehabilitation of style analysis as a precondition for a well developed musicology of performance.

I can illustrate some possible ways of addressing the first two of these problems with reference to a project recently initiated at the AHRC Research Centre for the History and Analysis of Recorded Music.[3] (We also have plans to address the third problem, but I will not discuss them here.) The project involves the development of semi-automated

---

[3] The project is a collaboration between Craig Sapp, Andrew Earis, and myself; details at www.charm.rhul.ac.uk/content/projects/chopin.html and www.mazurka.org.uk.

methods for the capture of accurate timing information from recordings of piano music—this is where most of the research effort has gone at the time of writing—and their application to a large corpus of recordings of Chopin's mazurkas. Since my focus in this paper is on approaches to performance analysis rather than substantive interpretation, I shall deal only with bars 1–16 from eight recordings by different pianists of op. 67, no. 4, one of them being the Rubinstein recording graphed in figure 3, but the methods I describe can be scaled up to much larger bodies of data.[4]

It is in the nature of a project of this kind that it is front-loaded. You have to make a significant investment in data acquisition, but when that is done, certain kinds of analysis become trivially easy: Figure 5, for example, shows tempo flexibility, like figure 1, but figure 5 is based on the relationship of each beat to the average, which makes it a musically more fine-grained measure than Bowen's. (While the horizontal axis represents date of recording, there is an argument that the pianist's date of birth is actually more revealing, on the grounds that most performers acquire their style of playing at a relatively young age.)
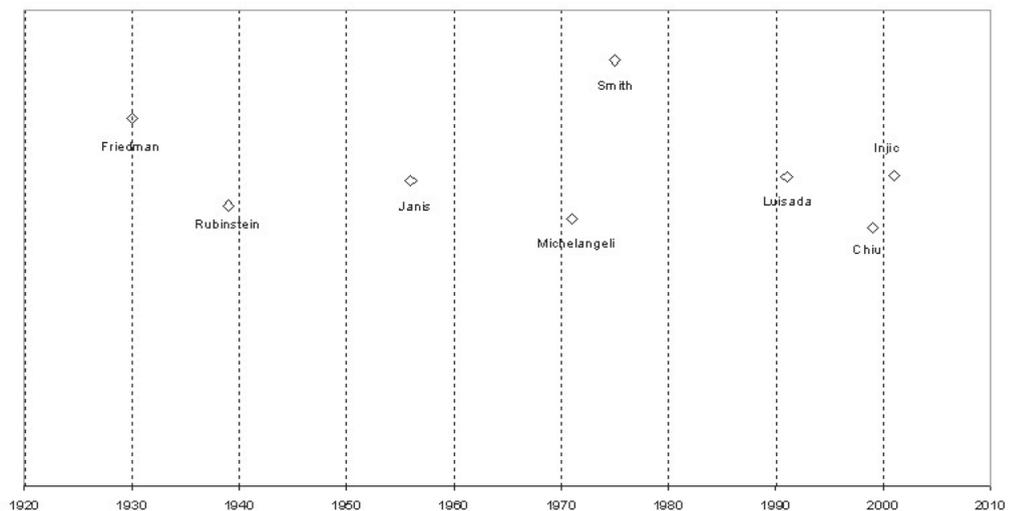
Fig. 5. Flexibility graph of recordings of Chopin's mazurka op. 67, no. 4, bars 1–16.

No correlational approach can tell you about the moment-to-moment decisions that lie at the heart of performance style, however, and so figure 6 is a version of figure 3 with the addition of data for the other seven recordings and an overall average. Graphs like this contain a great deal of information, but they do not communicate it effectively to readers. Perhaps a good way to think about this is to ask what this kind of graph actually represents. The obvious answer is that it represents the length of each beat, but a more telling one is that it represents a pattern of deviation as compared with strictly metronomic performance—the kind of "performance" for which a sequencer is required, since humans are incapable of performing without interpreting. (Such a metronomic "performance" would, of course, show up as a straight horizontal line.) To put it that

---

[4] The recordings (with original release dates) are as follows: Frederic Chiu, HMX 2907352.53 (1999); Ignazy Friedman, Philips 456 784–2 (1930); Eugen Indjic, Calliope 3321 (2001); Byron Janis, Philips 456 847–2 (1956); Jean-Marc Luisada, DG 463054–2 (1991); Arturo Benedetti Michelangeli, Philips 456 904–2 (1971); Artur Rubinstein, Naxos 8.110656–57 (1939); and Ronald Smith, EMI 724358576726 (1975).
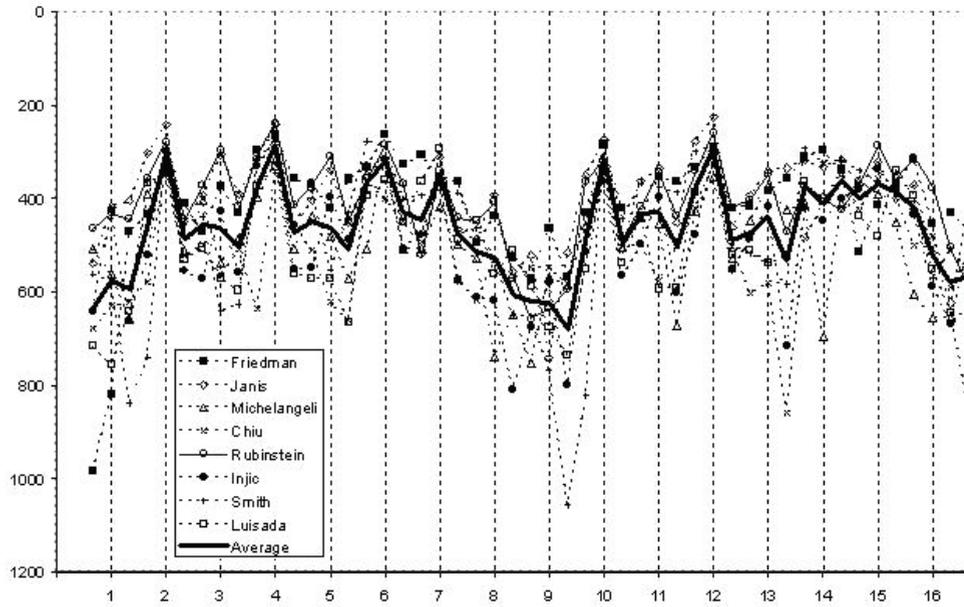
Fig. 6. Tempo graph of recordings of Chopin's mazurka op. 67, no. 4, bars 1–16, with average.

way raises the obvious question: What is the psychological reality of a metronome line, when nobody can play the music that way, and what leads us to think that it might be a useful yardstick for analysis?

The answer to the latter question is straightforward enough: Once again, it is the scriptist tendency to see performance as the reproduction—with certain deviations—of a text in which half notes are encoded as twice as long as quarter notes and quarters as twice as long as eighths. (Recall what I said about the objective time of twice as long and half as long being only a framework for negotiation.) Finding other ways to represent performers' shaping of tempo is less straightforward, though we are experimenting with a number of alternatives to the traditional graphic approach.[5] But even within the traditional graphic approach there are alternatives, some of which I would now like to briefly consider. The most obvious is to graph the pattern of divergence between tempo profiles that *do* have some kind of degree of psychological reality; the graph at the top of figure 7, for instance, measures Rubinstein's performance against the average profile from figure 6. (So the high values at the beginning of the upper graph reflect the fact that Rubinstein begins faster than average.) The rationale for this is that the average profile, which could, of course, be refined for a particular time and place, represents an aspect of the horizon of expectations against which an individual performance might be heard. That may or may not be a plausible argument, though there is perhaps evidence of psychological reality in the fact that synthesized performances based on average values generally receive good evaluations from listeners.[6] But even if it is not, the same

---

[5] Such as representations adapted from Craig Sapp's "keyscapes" (ccrma.stanford.edu/~craig/keyscape/); another example is the "performance worm" animations developed at the Austrian Research Institute for Artificial Intelligence (ÖFAI).

[6] Bruno Repp, 'The aesthetic quality of a quantitatively average music performance: two preliminary experiments. *Music perception*, 14/4 (1997) 419–44.
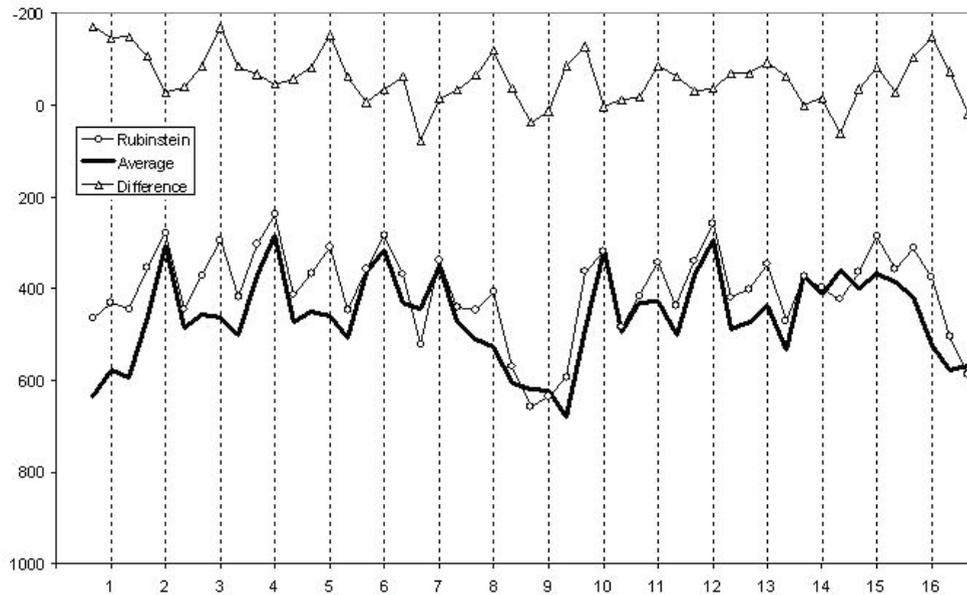
Fig. 7. Comparison between average tempo values in recordings of Chopin's
mazurka op. 67, no. 4, bars 1–16, and Rubinstein's 1939 recording.

method might be used to graph the profile of one performance against that of another, embodying the kind of relational listening of which I spoke earlier. Used this way, the graph becomes a diagnostic, directing your attention to those features of one performer's style that particularly distinguish it from another's.

A more principled way of achieving a similar end might be to stick with the conventional tempo profile but attempt to break it down into its components. I will illustrate what I mean by describing two such components, though more would be needed for the approach to be very useful (and here I am moving from things we have done to things we are planning). I have already referred to the characteristic way in which mazurkas are played, with the first beat the shortest, resulting in a kind of built-in syncopation or swing. One of the trivially easy analyses to which I referred focuses on how different performers interpret what I call the mazurka "script": Figure 8 shows the data for these eight recordings, and it can be seen that Rubinstein's version is quite close to the average (Friedman is the least typical, since his third beats are, on average, longer than his second beats). Of course, this begs the question of the extent to which performers play this script consistently—in other words, the extent to which the script is a constant or a context-dependent variable—but, by way of approximation, we could incorporate the average script values within a baseline for analysis of op. 67, no. 4. In figure 9 these script values have been subtracted from the values that make up the tempo profile: This removes at least some of the jitter in the profile attributable to the characteristic metrical shaping of the mazurka.

As for the second component, earlier I referred to the division of these 16 bars into two phrases of eight, and called it unsurprising—and this is more than a matter of common sense: It is well known from Neil Todd's work that performance of music of the so called common-practice style can be modelled as a series of superimposed arch-shaped profiles (at the four-bar level, eight-bar level and so on) expressed in terms
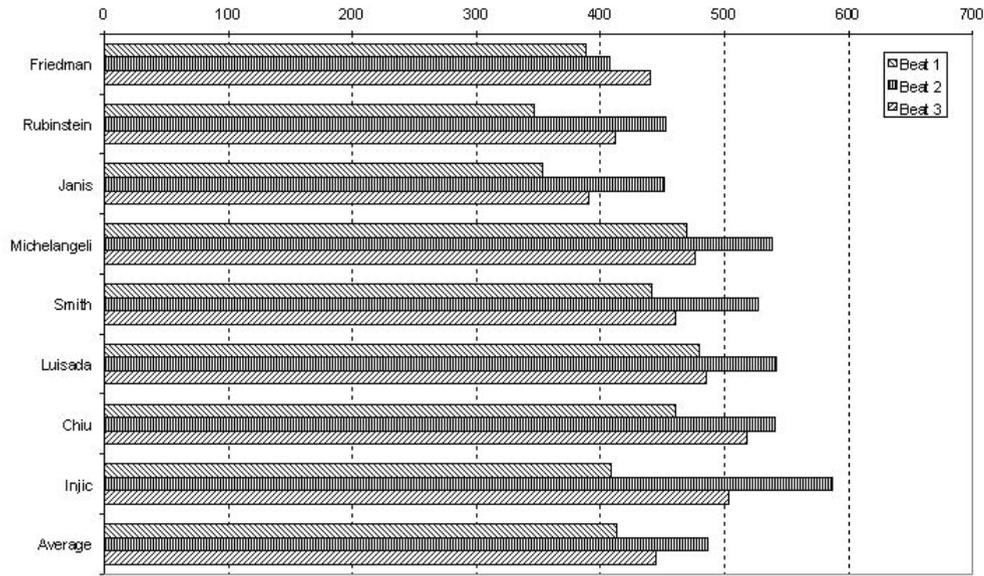
Fig. 8. Average duration of beats in recordings of Chopin's mazurka
op. 67, no. 4, bars 1–16 (milliseconds).

of both tempo and dynamics.[7] So we could treat this hierarchical phrasing model as a second component of a tempo baseline for op. 67, no. 4. And it is easy to think of other components that could be included within such an approach (some of the rules built into the expressive performance generation program *Director Musices*, by Anders Friberg, Johan Sundberg, and others, would be obvious candidates). The more such standard components of performance timing we can account for and subtract from the tempo profile of any individual performance, the more the resulting profile will highlight what is idiosyncratic about that performance rather than the general stylistic norms that inform it.

But we could go further than this. In terms of Todd's hierarchical phrasing model, what is striking about figure 9 is the absence of any clear articulation at the four-bar level. The trend line brings this out, and the note names below suggest a good musical reason for it: Despite the rhythmic repetitions and melodic sequences, the first eight-bar phrase is strongly based on a descending melodic arpeggiation of the tonic triad, and to this extent it functions as an irreducible unit. At least—and here is the point—that is how Rubinstein plays it. That means that we shall close in on one of the specific features of Rubinstein's playing of this piece, and be able to compare it more acutely with other performances, if we invoke not just the superimposed hierarchical phrasing model, but the relative weighting attached to different structural levels (high at the eight-bar level and low at the four-bar one). And at this point the approach becomes something more than a principled way of constructing a baseline tempo profile against which the idiosyncratic qualities of individual recordings stand out: It becomes a way of representing aspects of individual recordings. What I have been describing is, in fact, an application of a method recently developed by Peter Desain and Luke Windsor, among others, which not only models performances in terms of the relative weightings of a set

---

[7] First set out in Neil Todd, "A model of expressive timing in tonal music", *Music perception* 3/1 (fall 1985) 33–57.
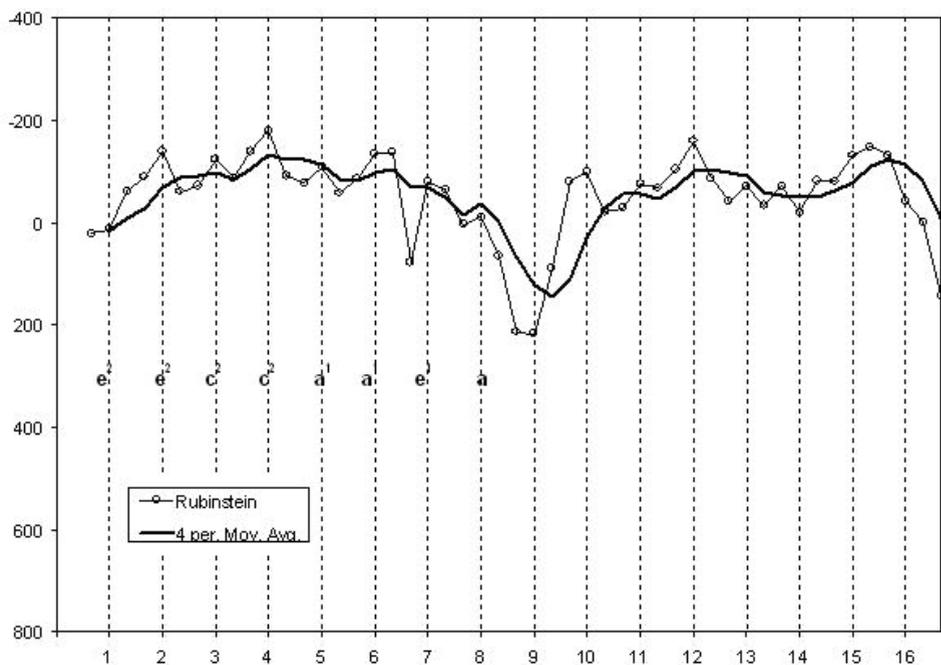
Fig. 9. Normalized tempo graphs of Chopin's mazurka op. 67, no. 4, bars 1–16
(average and Rubinstein 1939), with prominent melodic notes.

of predefined analytical components, but also derives the weightings mathematically from the performance timing data. (In principle the same is possible for dynamic data, but nobody has tried that yet.)[8]

One can then envisage a range of components and weightings—some of them, perhaps, based on score-based analytical approaches—which would together constitite a compact representation of performance style: a representation drawn (unlike Bowen's correlations) from the details of the performance's note-to-note progression, but at the same time (unlike the conventional tempo-profile approaches) enabling direct comparison across large numbers of recordings. We would have a means of characterising one performer's style in relation to other performers' styles across a number of common dimensions, as well as in terms of its idiosyncratic features. That, in turn, would provide a much better informed basis for the kind of correlational analysis exemplified by Bowen's graphs or figure 5 above; it would also open up possibilities for stylistic comparison across different mazurkas. In this way a common method could support performance analysis drawing on two distinct epistemologies: on the one hand refined perception and critical evaluation of individual performances (the equivalent of work-oriented structural analysis), and on the other hand more abstract and empirically determined mapping of trends in performance style (the equivalent of the largely defunct tradition of composition-oriented style analysis).[9]

----

[8] See Luke Windsor, Peter Desain, Amandine Penel, and Michiel Borkent, "A structurally guided method for the decomposition of expression in music performance", *Journal of the Acoustical Society of America* 119/2 (2006) 1182–93. My thanks to Luke Windsor for discussing this approach with me.

[9] I have explained what I see as the relationship between empirical method and epistemology at greater length in "Border crossings: A commentary on Henkjan Honing's 'On the growing role of observation, formalization and experimental method in musicology'", *Empirical musicology review* 1 (2006) 7–11.

But at this point I would like to address some of the possible problems with this kind of approach. Any computational style analysis poses the risk of being swamped by data that are too reduced to support musically interesting conclusions. It is easy to imagine objections along these lines: Do not beat durations need to be understood in terms of the specific rhythmic events that take place within the beats, for instance, and is durational information a sufficient basis for analysing performance style considering its interaction with other parameters such as dynamics or articulation?[10] (We hope to extend our data capture methods to include both sub-beat timing and dynamics, but the basic point is that all analysis involves data reduction.) It is also easy to imagine a more all-encompassing objection: Is not using computers to analyze recordings an example of analysis "turning ideas into objects, and putting objects in place of people",[11] so theorizing both performers and the experience of music out of performance?

This more all-encompassing criticism involves a false if widespread conflation of the use of computers and the retreat from experience, and as such merits a brief response. Recall where my discussion of computational modelling of performances came from: the attempt to establish a baseline for tempo analysis that has more psychological reality than the conventional metronome line. I suggested that one might instead focus on the patterns of similarities and differences between different performances, or between a given performance and some kind of customised baseline appropriate to the music in question, and so bring the individual qualities of the performance into high relief, before developing this into the idea that one might attempt to represent aspects of particular performances through a system of analytical weightings. And, as in the case of Anthony Pople's "Tonalities" software,[12] using the Desain/Windsor method means starting with a set of analytical parameters based on your intuitions about the performance in question, running the empirical analysis, checking the fit with the data, and using that as a basis for modifying the parameters and rerunning the analysis: It is an iterative process which can equally well be described as one of refining the fit between model and data, and of refining the analytical intuitions with which you began. In effect you are "tuning" your analytical experience of the performance, bringing it into the sharpest possible focus rather as you might adjust a microscope. The computer is being used, as is usually most productive in the arts and humanities, not as a substitute for the human researcher, but as an assistant, or, as Pople expressed it, "a junior partner … in an expert-to-expert dialogue about the analysis of specific pieces".[13]

But the objection about theorizing performers out of performance is not so easily disposed of, and leads to what I see as a general principle in performance analysis: no one method can capture more than one dimension of what is always a multidimensional practice, and therefore it is always necessary to combine methods. I can make this point by describing a specific case when empirical analysis would have been misleading if it had not been supplemented by reference to the performer. This comes from a project carried out by Eric Clarke, Bryn Harrison, and Philip Thomas as well as myself, which

---

[10] The classic discussion of such fundamental issues is Peter Desain and Henkjan Honing, "Tempo curves considered harmful", *Contemporary music review* 7/2 (1993) 123–38 (www.nici.kun.nl/mmm/papers/dh-93-f.pdf).

[11] Richard Taruskin, *Text and act: Essays on music and performance* (New York: Oxford University Press, 1995) 24.

[12] Anthony Pople, "Using complex set theory for tonal analysis: An introduction to the Tonalities Project", *Music analysis* 23/2–3 (July 2004) 153–94. See also www.nottingham.ac.uk/music/tonalities/.

[13] Anthony Pople, "Modelling musical structure", *Empirical musicology: Aims, methods, prospects*, ed. by Eric Clarke and Nicholas Cook (New York: Oxford University Press, 2005) 151.

involved Thomas learning and performing a piano piece composed by Harrison, with recordings being made of the rehearsals and the premiere, as well as a session at which they worked on the piece together.[14] The aspect of this project that is relevant here is our evaluation of the accuracy of Thomas's rhythmic interpretation.

Our first analysis treated the music as a single sequence of note attacks and correlated these with the specifications in the score; in terms of rhythmic detail the correlation was not very high, which might seem unsurprising given its complexity. (The passage with which I am concerned is notated in the score as three polyphonic lines—though that is not the effect of the music when you listen to it—with abundant irrational and sometimes nested rhythmic values.) But then we realized that the way we had analysed the music did not at all correspond with how Thomas had conceived it as he learned the piece. In Thomas's words,

> Each of the three lines has its own internal energy, which is difficult to sustain throughout its duration. I felt that the only way to learn this was to get to know each line on its own, and then in pairs before putting them all together, much like Bach counterpoint.… When putting everything together I would then find points that I could identify as markers, such as a downbeat in one voice which the other voices could respond to, always trying to hear the individual energy of each line as learned on its own.[15]

Reanalysed as three separate lines, it turned out that the pattern of note attacks was not just extremely close to the specifications in the score, but had been so almost from the very first time Thomas played through the music. Our initial evaluation of the rhythmic accuracy of Thomas's playing, then, was based on a misunderstanding of his performance strategy: We had interpreted the data in a manner that was mathematically right but musically wrong. And the point I am making is that it took a combination of different kinds of information—empirical and ethnographic data—to bring this to light. My own contribution to the project was an ethnographic study of what complex rhythms mean for the performer (and composer), based as much on what Thomas and Harrison said as on the performance data narrowly defined.

More generally, I would argue that the more you rely on hard, quantitative data such as inter-onset timings in analysing performance, the more important it is to triangulate this against ethnographic and other forms of qualitative data—the kinds of data familiar to popular musicologists, ethnomusicologists, and interdisciplinary performance theorists. (I am thinking primarily of theater and dance studies, but more generally, of approaches to performance as embodied action.) Of course, to take my argument back to its starting point, ethnographic approaches cannot be directly applied to historical recordings. But the point to make here is the one Peter Jeffrey made in relation to medieval chant: You maximize your chances of making sense of such meagre traces of past performance events as chant notations (and the same applies to recordings) by interpreting them in terms of a broad conception of performance that builds on every possible form of available evidence.

How, then, might it be possible to bring together the kind of empirical analysis I have been talking about and the broader issues of cultural meaning, on which I

---

[14] Eric F. Clarke, Nicholas Cook, Bryn Harrison, and Philip Thomas, "Interpretation and performance in Bryn Harrison's *être-temps*", *Musicae scientiae* 9/1 (spring 2005) 31–74.

[15] Ibid., 40.

touched near the beginning of this paper? It is too easy to assume that you need to do the empirical analysis before starting to think about cultural meaning: This strategy for analytical deferral (which, in effect, goes back to Hanslick) might be sidestepped by, for example, asking whether some of the differences in performance style emerging from the mazurkas project might not be explained in terms of the difference between rhetorical and structuralist approaches to performance. (Although rhetoric can be used to articulate structure—so, this is not a simple binary—rhetoric implies a reception-oriented conception of what music is, and the distinction between this and a structuralist conception takes one deep into the domains of ontology and aesthetics.) But I can briefly outline a more straightforward answer to this question with reference to a particularly telling example of changing performance style.

Webern's piano variations op. 27 were written just two years before World War II, and became well known only after the war, and specifically in the context of the Darmstadt avant-garde within which Boulez elevated Webern to the role of patron saint of new music. Op. 27 consequently took on something of the status of a sacred text, and was disseminated in the cool, abstract recordings of modernist performers such as Charles Rosen. But, as anybody knows who has read Peter Stadlen on this subject, Webern's own conception of the music was very different. Stadlen had given the premiere back in 1937, in Vienna, and had been intensively coached for the performance by Webern himself. As Stadlen records, Webern

> spent countless hours trying to convey to me every nuance of performance down to the finest detail. As he sang and shouted, waved his arms and stamped his feet in an attempt to bring out what he called the meaning of the music I was amazed to see him treat those few scrappy notes as if they were cascades of sound. He kept on referring to the melody which, he said, must be as telling as a spoken sentence. This melody would sometimes reside in the top notes of the right hand and then for some bars be divided between both left and right. It was shaped by an enormous amount of constant rubato and by a most unpredictable distribution of accents. But there were also definite changes of tempo every few bars to mark the beginning of a new "sentence".[16]

Although in 1979 Universal Edition published a facsimile edition of Stadlen's own score of op. 27, annotated with Webern's performance directions, none of what Stadlen describes can be traced in the standard published edition. So this is a classic demonstration of just how slender a basis scores can provide for musicological interpretation. However, it is equally a classic demonstration of the dangers of using recordings as a basis for musicological interpretation, for none of what Stadlen describes can be heard in the recordings released during the 1950s and 1960s either. Recordings like Rosen's, with their concern for objectivity and balance—you might say their structuralist rather than rhetorical orientation—construe op. 27 as an emblem of post-war modernism, which is to say modernism in the tradition of the Bauhaus and the "International" style. When cultural historians—historians of architecture as much as musicologists—speak of "modernism", that is usually what they mean. Yet with its focus on clarity of structural articulation, Bauhaus/International modernism is strikingly different from the Viennese variety, the central conception of which might be said to be the concealment of hidden meaning behind appearances. One of the musicological achievements of the last two

---

[16] Peter Stadlen, "Serialism reconsidered", *The score* 22 (1958) 12.

decades, made possible through exploration of the Webern *Nachlass* as well as increasing historical distance, has been the writing of Webern back into the context of pre-war Viennese modernism, with its strange juxtapositions of Neoplatonism, positivism, and Krausian ethics, and with the waltz as its ubiquitous soundtrack. And it is this reinvented, and surely more complex, image of Webern that sounds through such more recent recordings of op. 27 as, for instance, that of Mitsuko Uchida (2000).

The purpose of what I have said is not, of course, to stipulate how op. 27 should be played. It is rather to suggest that recordings of op. 27 embody changing thinking about Webern, which corresponds to the way in which the man and his music have been written into very different analytical and historical stories—stories that are animated by very different aesthetic, and perhaps also social and ideological, values. In that case, the kind of technical analysis that I have been discussing in this paper, tempo graphs and all, can become an instrument for refining and nuancing broad cultural concepts such as modernism: The music on the records—the authentic soundtrack of the 20th century—can function as a primary source for historical and critical understanding across the arts and beyond. In his article "Cinderella, or, Music and the human sciences", Leon Botstein has called for music to be enlisted "as a primary vehicle for the reinterpretation of culture and society".[17] I would like to think that the kind of performance analysis I have sketched in this paper could be a step towards doing just that. And if that is right, then writing performance into the mainstream of musicology may be the key to completing the job that the "New" musicologists began.

---

[17] Leon Botstein, "Cinderella, or, Music and the human sciences: Unfootnoted musings from the margins", *Current musicology* 53 (2002) 134.