9 Methods for analysing recordings

NICHOLAS COOK

If analysis means studying something in order to gain knowledge and understanding of it, then there are any number of ways of analysing recordings, and any number of reasons for doing so. Performers, recording engineers, historians of recording technology and historians of performance practice listen to recordings with quite different kinds of knowledge and understanding in mind: analysis means different things to them. The same applies to acoustic scientists, record collectors and archivists, or communication theorists, not to mention people in the A&R divisions of record companies whose job is to spot the next big hit. The list goes on.

This chapter basically assumes that your reason for analysing recordings is to gain a better understanding of them as culturally meaningful objects, and more specifically that you are primarily interested in the effect of music as experienced in performance, whether live or recorded. In that sense its orientation is musicological, although that too is a term that can be defined in different ways. Recordings are a largely untapped resource for the writing of music history, the focus of which has up to now been overwhelmingly on scores, and recent technological developments have opened up new ways of working with recordings – ways that make it much easier than before to manipulate them, in the sense that we are used to manipulating books and other written sources. I begin by introducing software that makes it possible to navigate a number of different recordings, and to create visualisations that help to heighten aural understanding of what is going on in the music. (Actually such software could be useful for practically all the people I mentioned in the first paragraph.) I move on to approaches that involve the comparison of large numbers of recordings in order to identify and characterise stylistic elements. Such approaches might be described as musicological in a relatively narrow sense. But then, in the final section, I consider some critiques of such approaches – critiques that have come from both within and outside musicology – and set them into the context of more broadly cultural approaches to recorded music.

Extending the ear

Important musicological work has been carried out using equipment no more specialised than a record or CD player, a pencil, and perhaps a stopwatch, coupled with the capacity for close listening that comes with experience. An example is the work of Robert Philip, whose two books between them represent a first draft of the history of classical music performance during the twentieth century. The quantitative dimension of his research hardly goes further than tables of performed tempi at various points in different recordings, while the quality of his listening is captured in passages such as that describing Ignaz Jan Paderewski's 1930 recording of Chopin's Mazurka Op. 63 No. 3:

> At the beginning he establishes a rhythm with a long first beat and a short second beat … This is varied at points of particular emphasis. For example, the approach to the highest point of the melody at bars 4–5 is emphasized by shifting the tenuto to the second beat … At bar 5 the lengthening of the second beat underlines the start of the phrase, and further emphasis is given by arpeggiating the accompaniment and delaying the melody note.[1]

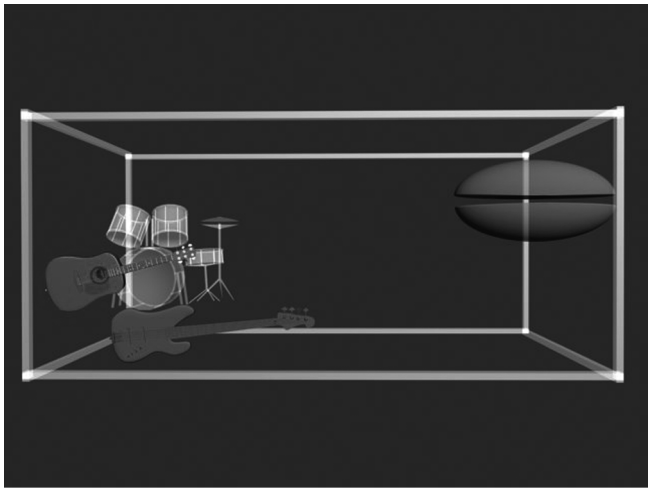
Much the same might be said of the account of Jimi Hendrix's Woodstock performance of 'Star Spangled Banner' that Eric Clarke offers in his book *Ways of Listening*, which is not based on the empirical approaches through which Clarke made his reputation, but relies on straightforward verbal description. In the course of an argument that Hendrix's adaptation of the American national anthem derives much of its meaning from the clash between 'official' culture and (then) counter-culture, Clarke observes that the G♯ eight seconds into the performance (the highest note of the opening arpeggio) is 'approached by a small but clearly audible pitch bend or glide up to the note from the preceding E4 – a characteristic stylistic invariant for rock-guitar playing'.[2] He cites some more examples, and then concludes, 'At the same time as the anthem is specified by its intervallic and rhythmic invariants, rock as a genre is specified by invariants of performance. The cultural clash is directly specified in the material itself.' To be sure, my characterisation of this account as 'straightforward' may have been misleading, to the extent that Clarke's purpose is in part to set out an approach informed by the ecological psychology of J. J. Gibson. But however sophisticated the theoretical approach, it is grounded in the act of listening. For musicologists at least, that is where all analysis of recordings must start.

It is however possible to use new technology to create an environment that makes it easier to listen effectively, in the sense of moving around a recording to compare different parts of it, or moving between different

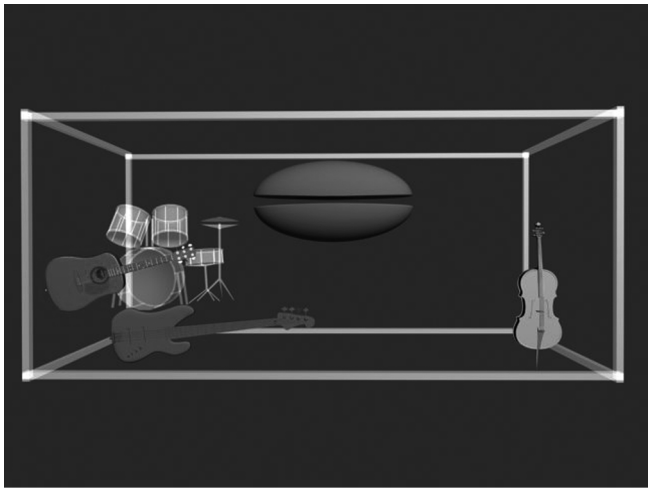
recordings to hear one against another. In this chapter I demonstrate such possibilities through the use of Sonic Visualiser, a free program developed at Queen Mary, University of London, but some at least of its functionality is available in other programs, or is likely to be in the future: I don't provide detailed instructions on the use of Sonic Visualiser here, but they are available in web-based tutorials designed to complement this chapter.[3] In addition to the familiar wave-form representation and playback controls shown in Figure 9.1, Sonic Visualiser provides two features that are particularly powerful for working with recordings. One is the ability to annotate the sound file, for instance by marking where each bar occurs: you can tap to the music as you listen to it, and use the resulting barlines to navigate the recording. (You can see the barlines in Figure 9.1.) The other is the ability to align multiple recordings of the same piece: Sonic Visualiser will work out which point in one sound file corresponds to the same point in others, so that you can – for example – go straight to bar 9 of each. The importance of these apparently simple features should not be underestimated. They create the same kind of environment for recordings that is taken for granted when working with scores or other written documents, where you can flick back from one page to an earlier one, or place several scores side by side to compare them. The effect is to give a new dimension to close listening.

As its name implies, Sonic Visualiser also offers a range of features for visualising what you hear, but before I discuss these it's worth considering what is gained by visualising music. Actually visualisation is a fundamental analytical technique: established score-based analytical methods employ a wide range of notational or graphical representations that sometimes help to bring what you hear into focus, and in other cases complement what is readily audible (schematic representations of sonata form or the 12-bar blues, for instance, make it easier to perceive the pattern in what may sound like a mass of details). Of course, traditional printed images are less compelling than animations that move in time to the music, as anybody knows who uses Windows Media Player. And while Windows Media Player visualisations are not designed to focus attention on the music in an analytical sense, there are other ways of visualising music that do just that. Figure 9.2, for example, shows how the voices (shown by lips) and instruments are located within the stereo sound space during the opening verse (from 0′30″) and after the chorus (from 1′09″) of 'King Midas in Reverse' by the Hollies; in the original animation the image changes as the sound sources do. This visualisation represents something that is there to be heard in the music, but it adds something to the experience, refining and focusing your listening, and making you more aware of the sound space. It is in this sense analytical, and on the basis of a number of such analyses Ruth

Figure 9.1 Working with multiple files in Sonic Visualiser



(a)



(b)

Figure 9.2 Soundbox images of the Hollies, 'King Midas in reverse'

Dockwray and Allan Moore, who developed this form of representation, have created a taxonomy of sound spaces and initiated historical interpretation of a previously undocumented aspect of recorded music practice.[4]

Figure 9.2 is computer-generated, but only in the sense that that is how the graphics have been created: it embodies the outcomes of close listening. Programmes like Sonic Visualiser, however, can generate a range of visualisations directly from the audio, and some of these are also effective in refining and focusing the listening experience. The most flexible of these visualisations are spectrograms, which represent sound in three dimensions: time (from left to right), frequency (from top to bottom), and intensity (by means of colour or, in black and white, shading).

Figure 9.3a is a very zoomed-in image of the passage from Hendrix's 'Star Spangled Banner' that Clarke discussed, showing just the fundamental frequency: the pitch bend, now seen in great detail, lasts from 8.0″ to 8.2″. More typical – less zoomed-in – spectrograms show pitches as several parallel lines because they include not just the fundamental (the frequency that corresponds to the pitch we usually hear) but also the individual harmonics at integer multiples of the fundamental: as an illustration, Figure 9.3b shows bars 42–6 from Sophie Braslau's 1928 recording of Schubert's song 'Die junge Nonne'. (The relative strength of the harmonics is important because it is one of the determinants of tone quality.) The sawtooth-like waves in Figure 9.3b represent Braslau's vibrato; you could easily measure its speed or depth if you wished. In the lower part of the image, and quite distinct, are the piano notes: because there is no vibrato they show up as straight lines, sometimes with an initial wedge shape resulting from their dynamic profile (sharp attack followed by decay). As Figure 9.3 demonstrates, spectrograms can be quite variable in appearance, because there are many different settings which enable you to focus on particular aspects of the sound at the expense of others, and there are also different colour schemes. But they all represent sounds using the same three dimensions, and are therefore read in the same basic manner.

Conventional score notation is extremely selective as a representation of musical sound: it provides a basic pitch and time framework with some annotations, but gives only broad indications regarding dynamics, articulation and timbre, and says virtually nothing about temporal or dynamic nuance. By being so selective, it can convey those aspects of music that it *does* convey very clearly. Spectrograms are just the other way round. Their attraction is that in principle all aspects of the sound are present in them; the downside is that in practice it may be hard to extract the information you want. They are most useful for homing in on the details of performance – the unnotated nuances that are responsible for so much of music's meaning – and it is in this role that they have been used by such musicologists as Robert Cogan or Peter Johnson in the field of classical music, and David Brackett or Serge Lacasse in popular music. When they are integrated into the working environment for studying recordings, as in Sonic Visualiser, they help to transform listening into analytical interpretation. Figure 9.3b comes from an article by Daniel Leech-Wilkinson which compares a number of recordings of 'Die junge Nonne' in order to show how different singers shape their performances so as to imbue the song with quite different expressive meanings. He writes of the passage shown in Figure 9.3 that the 'very fast (0.03 to 0.05 seconds) swoops up to notes, which until now have made Braslau sound dramatic (or matronly, if you prefer), suddenly become swoops down from above, sounds rarely

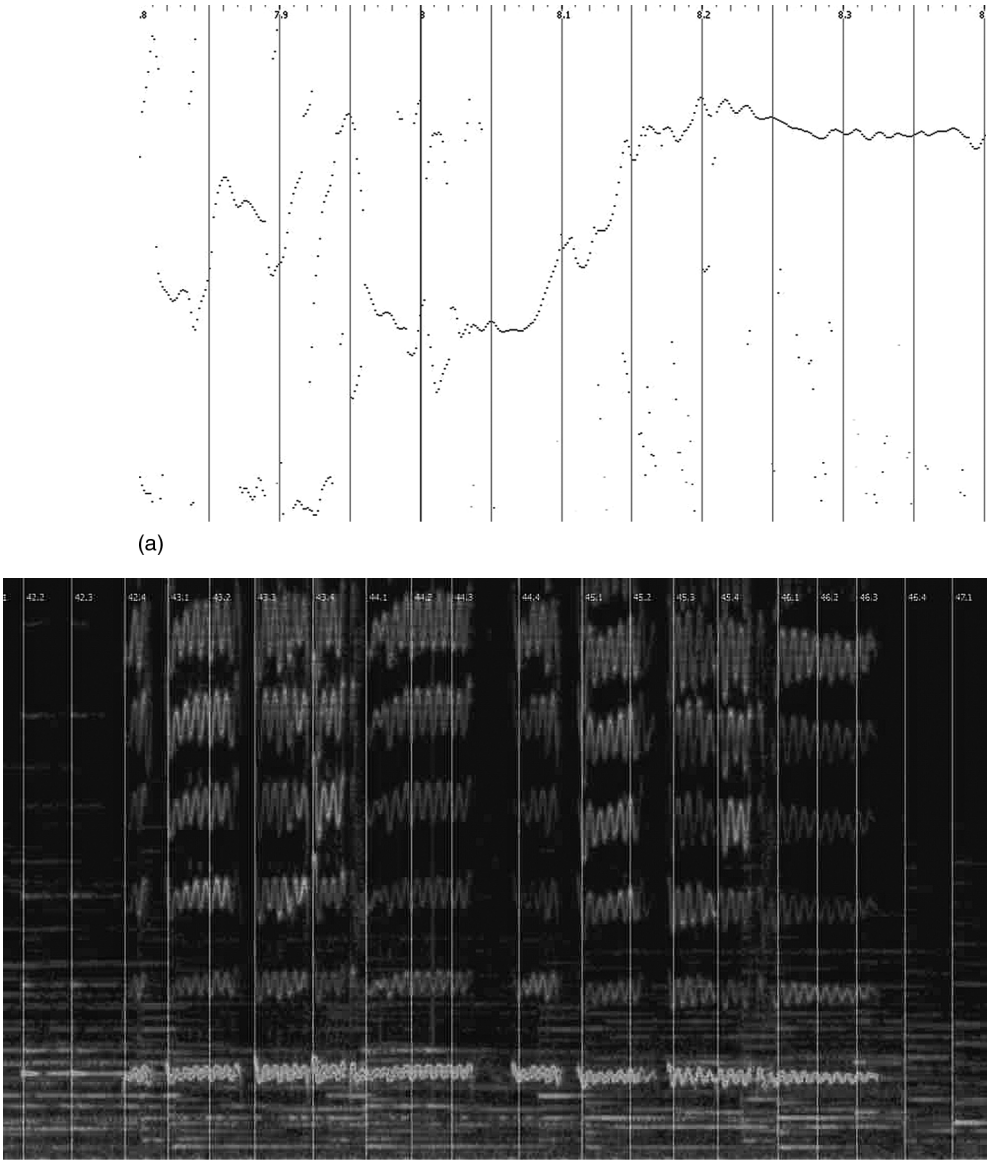


Figure 9.3 Spectrograms from (a) Jimi Hendrix's Woodstock recording of 'Star Spangled Banner' and (b) Sophie Braslau's recording of Schubert, 'Die junge Nonne'

used in song except in the Italian tenor sob: here also the start of each note is suggestive of crying'.[5] (The swoops from above are most easily seen in the lowest vocal line – the fundamental – for example on the first and third beats of bars 43 and 45: Leech-Wilkinson is arguing that inter-war recordings like Braslau's emphasise the dimensions of fear, horror and death, in contrast to later, less doom-laden interpretations.) Of course such effects might have been described without the use of a spectrogram, but it would

have been harder to be sure exactly what aspects of the sound are responsible for them, or to communicate them to readers.

If a limitation of spectrograms is that it can be hard to extract the information that you want from them, then an alternative approach is to extract just those aspects of the sound you are interested in and create customised ways of representing or manipulating them for analytical purposes. Typically such aspects include timing and dynamic information. Though work of this kind has become increasingly common in musicology over the last two decades, the methods were principally developed in psychology and cognitive science. The extraction of timing information from recorded sound goes back to Carl Seashore's work in the 1930s, but the modern foundations of this approach lie in a series of articles published by Bruno Repp in the late 1980s and 1990s: Repp used a waveform editor (nowadays Sonic Visualiser could be used for this purpose) to locate the beginning of each note by eye and measure the time interval between notes. This visual approach was very laborious but yielded a representation of the temporal profile of each recording sufficiently accurate to support Repp's detailed analyses of the data.

A representative example is his 1992 study of twenty-eight recordings by well-known pianists of Schumann's 'Träumerei',[6] in which the analytical results might be summarised under three headings. First, virtually all the pianists marked the large structural divisions of the music by slowing down at the end of sections: there was little significant variation between pianists at this level. Next, Repp carried out a form of factor analysis on the timing profiles *within* these sections: this is a statistical technique that reduces the complexity of large data sets by extracting the principal components. If there was basically only one way of playing the piece, with a certain amount of semi-random variation, then the analysis would yield only one factor. In fact the analysis yielded three factors, one of which was shared by a large number of pianists and the other two of which were respectively associated with Horowitz and Cortot: Repp saw these as representing distinct interpretive strategies, elements of which might to some extent be mixed in specific performances. The final element of Repp's analysis consisted of extracting the timing data for the most striking melodic gestures in the music, and fitting them to mathematical functions: he found that parabolas generally yielded the best fit, which suggests that the practice of slowing down at the end of a melodic gesture may form some sort of correlate of the motion of objects in the physical world (if you throw a ball into the air, its arc will describe a parabola). Repp found that virtually all the performances exhibited these parabolic functions, but with significant differences of scale between different performers.

Repp's articles constitute a storehouse of analytical methods which musicologists have perhaps not sufficiently explored. At the same time, as a psychologist, Repp was more concerned with discovering general principles underlying the distribution of the data than trying to engage with the aesthetic properties of specific performances, and he was not concerned at all with issues of cultural meaning. This is a way of saying that not all analysis of recordings is musicological in intent, and the same applies to cognitive-scientific approaches, of which the outstanding example is perhaps the work of Gerhard Widmer and his co-workers at the Austrian Research Institute for Artificial Intelligence. Rather than using Repp's visual method to extract the basic data, this research was based on a semi-automatic system which extracted both the timing of the beats and their associated dynamic values (the system was semi-automatic because a fully automated system could not achieve the necessary accuracy, so that it was necessary to edit the data manually). The resulting tempo and dynamic data were then input to a visualisation system called the 'Performance Worm': this is a computer animation that moves as the music plays, with tempo on the horizontal axis and dynamics on the vertical axis. Figure 9.4 is the image generated by the first four bars of Daniel Barenboim's recording of the second movement of Mozart's

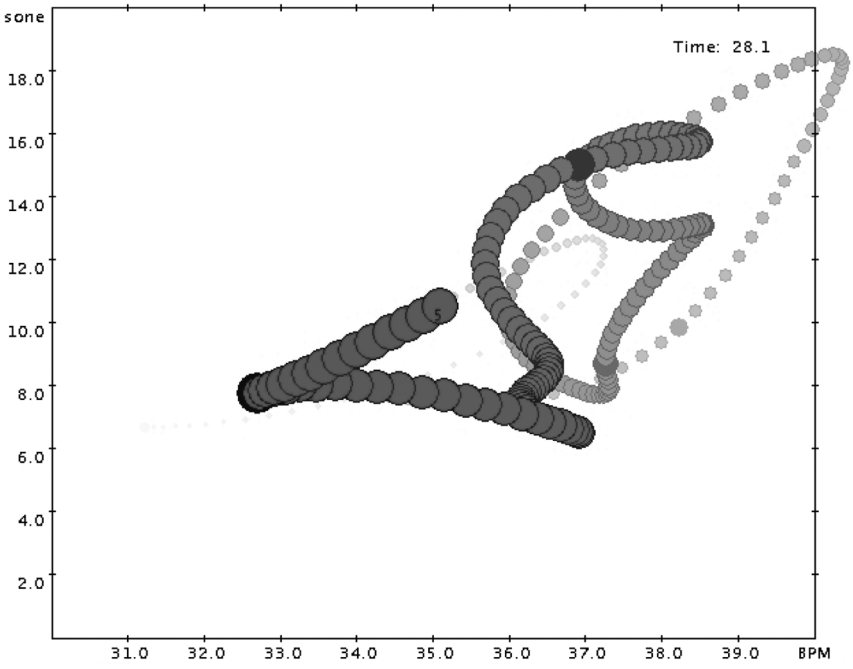


Figure 9.4 'Performance Worm' visualisation of Mozart's K. 332, bars 1–4, performed by Daniel Barenboim

Sonata K. 332,[7] with the darker sections representing the worm's most recent movements: the trails fade with time.

But while this is an intriguing way of visualising performance and bringing out certain of its gestural qualities, it was only the first step in a more elaborate analytical process of which I can give only a bare summary. The complete worm trails for four pianists' recordings of five Mozart sonatas were divided into short segments and subjected to cluster analysis, yielding an 'alphabet' of prototypical tempo–dynamic patterns as found in these performances. These prototypes were then organized into a matrix of 8 cells by 5, with a self-organising map algorithm being used to place similar cells adjacent to one another. The resulting images for each of the four pianists are shown in Figure 9.5, where the map-like shading shows how frequently the prototypical patterns were found in each pianist's performance. (Lighter shade means higher frequency.) It is obvious that the pattern for each pianist is quite distinct, with Pires perhaps being the most idiosyncratic.

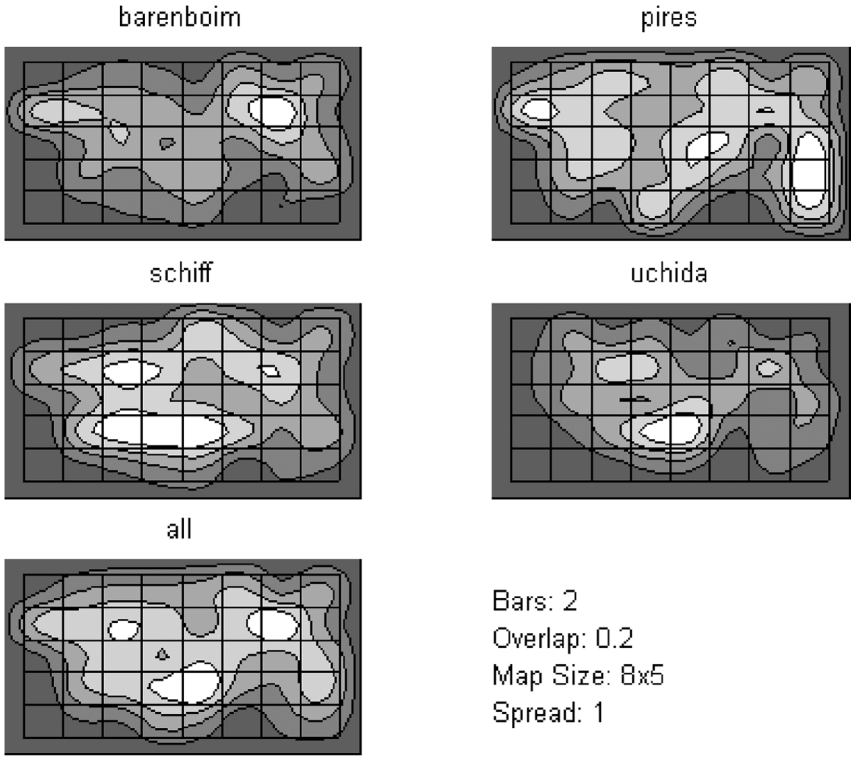


Figure 9.5 Images of four pianists' performances of five Mozart sonatas plus average values, based on tempo–dynamic associations

For the musicologist, work like this raises questions such as whether it is possible to make musical sense of the stylistic characterisations in Figure 9.5, or whether the analysis is at too abstract a level to be brought to bear on musicological issues such as aesthetic effect and cultural meaning. To raise such concerns is not to criticise Widmer's project, which is primarily a study in artificial intelligence rather than musicology: that is, it is a highly impressive attempt to model aspects of a particularly complex human behaviour – piano performance – through a range of objective methods. While musicologists have also used visualisations based on abstracting timing and sometimes dynamic information from the sound, they have done so with different purposes and, it has to be said, in general with much less technical sophistication. During the 1990s there was a considerable amount of work based on a tapping approach: you listened to the music, and tapped on a computer to mark certain points (usually bars or beats). The computer logged the times at which you tapped, and this information was imported into a spreadsheet, with the normal output being a tempo graph. Taken from my 1995 article on two recordings by Wilhelm Furtwängler of the first movement of Beethoven's Ninth Symphony,[8] Figure 9.6 was generated by tapping once a bar, with the plotted values being the average of three separate passes.

This method was relatively quick and easy, but limited in accuracy – partly because of problems in coordination between ear and hand, and partly because the resulting data were a mix of anticipation and reaction. (Really you were measuring not the music, but your own

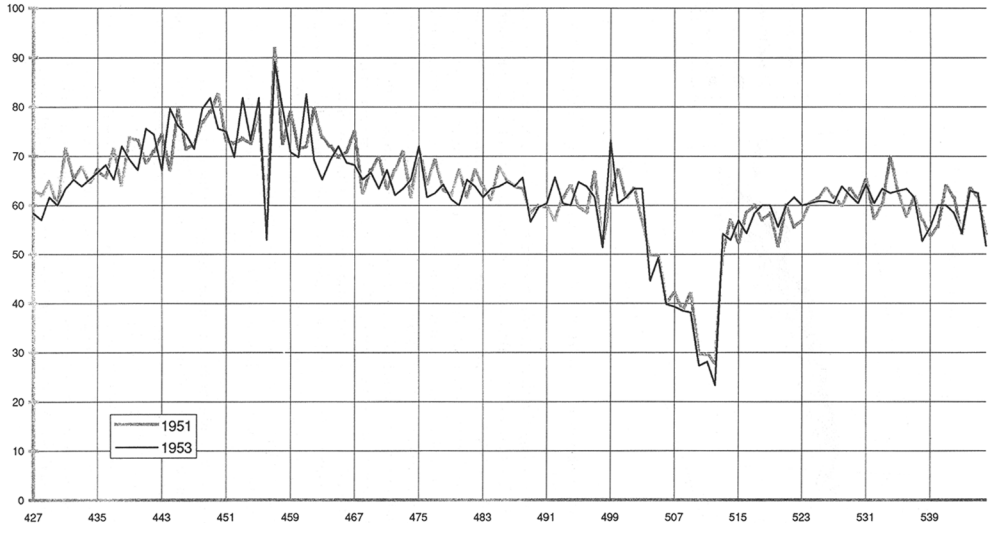


Figure 9.6 Tempo graph of Furtwängler's 1951 and 1953 recordings of Beethoven, Symphony No. 9, coda

physical response to it.) The data were probably good enough if you were tapping once a bar, but perhaps not if you were tapping once a beat. Nowadays, however, tapping can be carried out in an environment such as Sonic Visualiser, which offers crucial advantages: you can tap the beats and then listen to them as you play back the music, and you can then edit them, if necessary slowing down the playback, until you are confident they are where you want them. There are also plugins for Sonic Visualiser which make the onsets stand out visually in the waveform, while programs are being developed that take the tapped beats as their input and generate more accurate timing data, not only for beats but also for other onsets, along with associated dynamic data. While the data can then be analysed mathematically, as in Repp's and Widmer's work, tempo and dynamic graphs or other representations can be used together with spectrograms to create an integrated environment for working with recordings in which sound is combined with several complementary visualisations, each designed to bring out a particular aspect of the performance (see Figure 9.7). The result of all this is that the preparation of tempo or dynamic graphs no longer marks the end of the process, as all too often seemed to be the case in the past: nowadays it means you are ready to start on the real work of analysis.

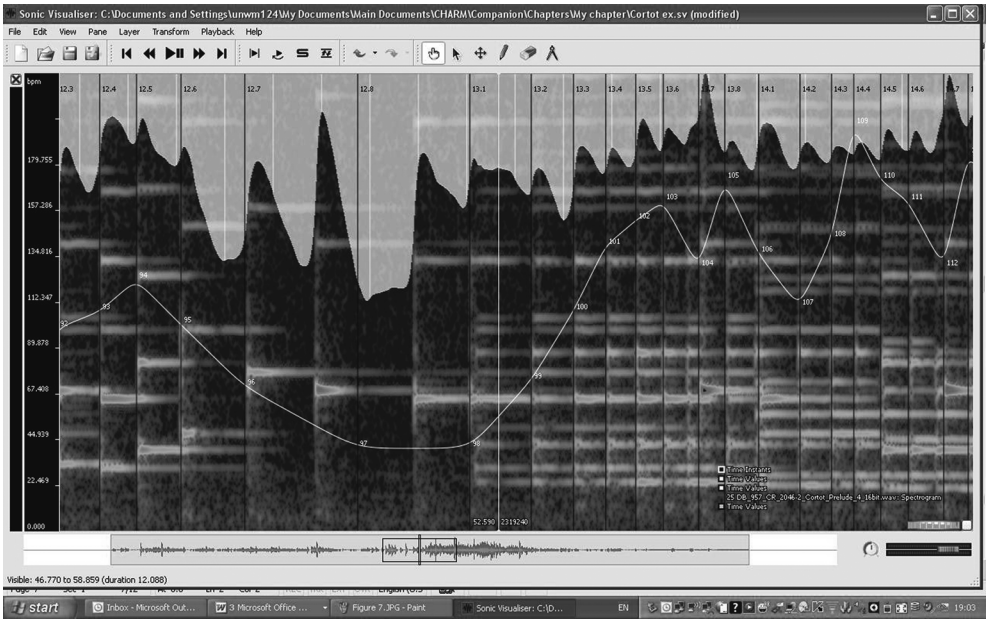


Figure 9.7 Using Sonic Visualiser to work with Chopin's *Prélude*, Op. 28 No. 4 in E Minor, bars 12–14, as recorded by Cortot in 1928 (Daniel Leech-Wilkinson). The display consists of a spectrogram, tempo graph, 'silhouette' representation of dynamics, and bar:beat numbers (measured in quavers). The main structural division of the piece falls at bar 13:1

An ear for style

The technical problems that constrained musicological analysis of tempo and dynamics are, then, being solved. Perhaps more intractable are the issues of interpretation involved in such work. The purpose of the 1995 study from which Figure 9.6 is taken was to investigate how far Furtwängler's performances could be understood in terms of Heinrich Schenker's analysis in his 1912 monograph on the Ninth Symphony. The underlying assumption was that what Schenker saw as discrete structural units would correspond to continuous tempo profiles – usually arch-shaped profiles – in Furtwängler's performance, with the breaks between sections being marked by rallentandi or caesurae. In this way the basic strategy was to begin with Schenker's analysis, and to see how far it could be mapped onto the performance. This, it seemed to me, was a valid approach because it is what Furtwängler himself must have done: he read Schenker's monograph shortly after it was published and was so impressed by it that he sought Schenker out, and the two men maintained a friendship until Schenker's death in 1935. More particularly, Furtwängler is known to have discussed the repertory he conducted with Schenker. The article, then, traced the musical consequences of a historical relationship.

But that was a special case. The problem occurs when the same approach is used in the absence of such a historical relationship, which can only make sense if one assumes that analytical approaches such as Schenker's embody fundamental musical principles that inform performances by artists who have never even heard of Schenker. That is a very large assumption to be making in the early stages of what is still a relatively new field of study. And the practice of working *from* a score-based analysis *to* a recording basically declares off limits all those aspects of performance that cannot be directly related to notational categories; it eliminates most of what there is to study before you even start, including all the rhetorical, persuasive, or expressive effects that contribute so much to the meaning of music as performance yet have little or nothing to do with structure as the music theorist sees it. (I note with embarrassment a cheap jibe against Mengelberg in my 1995 article, whom I described as 'a rubato conductor, a virtuoso', in other words, not a structural performer.[9]) Finally, to work from page to stage, as they say in theatre studies, is to treat a performance as first and foremost a reproduction of the musical work as embodied in the score; nobody would wish to say that that is *all* a performance is, but once you have started down this road, it is very hard to do justice to the creative dimension that makes it worth studying performance in the first place. And of course page to stage approaches simply do not apply to most music outside the Western classical tradition.

There is one further criticism I would direct at traditional musicological analysis of performance, recorded or otherwise. I mentioned the underlying principles regarding the relationship between score-based analysis and performance on which my 1995 article was based, but they are neither well developed nor explicitly set out. This failing is quite general among musicologists,[10] and the lack of well-articulated principles for the mapping between analytical and performance data results in a rather loose discourse in which tempo or dynamic graphs may not really provide the empirical support that is ascribed to them. There is also a tendency to see tempo profiles as objects of analysis in their own right, whereas according to Henkjan Honing and Peter Desain the tempo curve (as they term it) 'lulls its users into the false impression that it has a musical and psychological reality. There is no abstract tempo curve in the music nor is there a mental tempo curve in the head of a performer or listener.'[11] It is hard to be sure what to make of this argument: tempo is obviously linked to the ebb and flow of what we experience when we listen to music. But even so, Honing and Desain would claim that the tempo profile results from an indefinite number of different factors which really need to be understood individually. They propose that relevant factors might include the 'composers' pulse' patterns associated with the work of Manfred Clynes, the hierarchical phrase arching associated with Neil Todd, and the rule-based model of expressive performance developed by Johan Sundberg and his co-workers: to make sense of a tempo profile, then, it is necessary to break it down or 'decompose' it into its various components.[12] Whether these particular factors are necessarily the right ones is a matter for debate (especially since there is considerable overlap between Clynes's, Todd's and Sundberg's models), but the principle is persuasive.

Taken together, these criticisms suggest some profitable directions for musicological analysis of recordings. In the first place, both the availability of more robust data, and the possibilities afforded by programs like Sonic Visualiser of integrating them into the study environment, should encourage work that is grounded in close observation of recordings and builds towards appropriate analytical models, rather than importing its models wholesale from score-based analysis. Second, whereas both the CD-and-pencil and early tapping approaches encouraged work based on perhaps just one or two recordings analysed in relation to the score – as in the case of my Furtwängler study – new technology makes it much easier to draw comparisons between large numbers of different recordings, whether by means of Sonic Visualiser's alignment facility or the data extraction approach pioneered by Repp: this is tantamount to a shift of emphasis from the work (understood in relation to the score) to performance style, understood

through comparisons between recordings. The final element would be a move towards the identification of significant features that underlie tempo or dynamic profiles, and when I say 'significant' I mean to imply a need to consider the role that such features play in the communication of structural, expressive, or connotational information through performance.

There is a good deal of work that illustrates the first two of these directions. José Bowen and Eric Grunin,[13] among others, have made use of scattergrams that map simple performance features such as average tempo or duration (not necessarily the same thing, owing to repeats) against date of recording or performer's date of birth; their use of large numbers of recordings bolsters confidence that the resulting distributions are statistically significant.[14] Figure 9.8 is a rather more sophisticated example, showing how performances of the exposition from Beethoven's Third Symphony have in general become less flexible over some ninety years of recordings. Those by Furtwängler have been picked out, and are rather consistent in their degree of flexibility. (To measure flexibility, the exposition was divided into twelve sections and the relationship between their average tempi calculated.) Again, Richard Turner has used clustering software to group recordings of Brahms's First Symphony according to the similarity of their tempo profiles.[15] But although such analyses are based on direct comparisons between different recordings, rather than referencing them to the work they are performances of, they actually tell us little

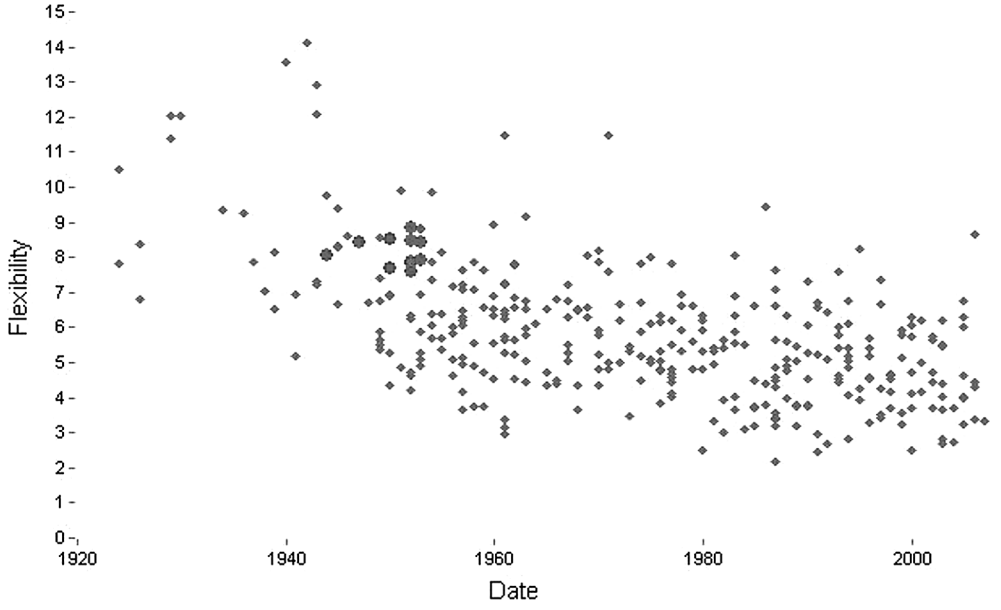


Figure 9.8 A comparative measure of flexibility in recordings of Beethoven's Symphony No. 3 plotted against date of recording, with Furtwängler's recordings highlighted

about performance style. The reason is obvious: not only are they exclusively based on tempo data – just one aspect of the performance, though an important one – but they also reduce the temporal evolution of the music to a single value, and in this way conflate quite different things. (A performance that swings wildly between frenetic tempi and funereal pauses may end up with exactly the same average tempo as one in the post-war sewing-machine style.) They are also heuristically unproductive. By this I mean that they are hard to relate to the music as experienced, and hence not effective in directing attention to specific points in the music that might reward further study. The danger is that they may close down rather than open up further investigation.

Other recent approaches retain the temporal dimension and so overcome some of these problems. Craig Sapp's multicorrelational plot of Artur Rubinstein's 1966 recording of Chopin's Mazurka Op. 30 No. 2 is shown in Figure 9.9a.[16] It is based on tempo data (like the Grunin and Turner analyses), and shows which other recording of the same piece is most similar at any given point; the various shades of grey are keyed to the other thirty-two recordings in this sample. What the plot is saying is that – despite Rubinstein's reputation as a performer who successively reinvented himself – much the closest match is with his own 1952 recording, with his 1939 recording also being closer than anyone else's (though not as close as the 1952 one). Here, however, it is possible to locate the particular points at which other recordings are most similar, and in this way gain some insight into what underlies the overall correlations. The similarity with Chiu's 1999 recording, for example, is confined to one particular point about three quarters of the way through (it is represented by the diamond-like shape), so you could listen to this passage in the two recordings and assess how far the connection makes musical sense. In this way a visualisation based on objective measurement can act as a prompt to further critical study, sending you back to the recordings with specific questions in mind.

Sapp's visualisations focus on style, in the sense that they are based wholly on comparison, and are resolutely bottom-up. But they are vulnerable to the Honing–Desain critique, in that the analysis is based on the overall tempo profile without any attempt to distinguish the different features contributing to it. The problem is one that Repp encountered in his 'Träumerei' analysis: he initially carried out factor analysis on the complete timing data, but only one component emerged, as the data were swamped by the slowing down at the end of major sections that was a feature of virtually all the performances – which is why he then carried out his factor analysis *within* these sections.[17] Sapp has addressed this problem through alternative visualisations based on smoothed and residual data: you smooth the original data mathematically, which brings out larger features such as phrasing, and then you subtract the smoothed

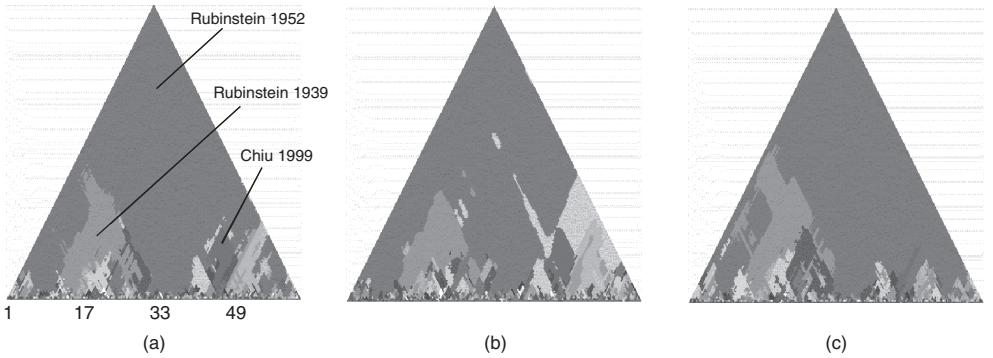


Figure 9.9 Multicorrelational tempo plots of Rubinstein's 1966 recording of Chopin's Mazurka, Op. 30 No. 2, using (a) full, (b) smoothed and (c) residual tempo data (Craig Sapp). Numbers represent bars. The base of the triangle represents the moment-to-moment succession of the music; the vertical dimension shows similarities at successively higher levels

data from the original data, so eliminating the large-scale features that swamped Repp's results and in this way focusing on smaller-scale features such as accentuation. Figures 9.9b and 9.9c show the smoothed and residual equivalents of 9.9a, and as can be seen the correlations change quite noticeably: the similarity to Chiu virtually disappears in the residual data, so must relate to larger-scale features. Sapp also uses similar full, smoothed, and residual plots based on global dynamics, and on the combination of tempo and dynamics.

In a sense, both Repp's distinction between what happens at the level of the whole piece, within sections, and in terms of melodic gestures, and Sapp's use of smoothed and residual data represent attempts to decompose the overall tempo into distinct features, in the way that Honing and Desain called for. But such approaches are perhaps most musicologically interesting when they correspond to features of the performance that make immediate musical sense, such as the hierarchical phrase arching that Honing and Desain put forward as a candidate for decomposition. As I mentioned, such phrase arching is associated with the work of Neil Todd, who developed a model of expressive performance based on the idea that performers give temporal and dynamic shaping to musical phrases through the use of parabolic functions (there is a link with Repp's work on melodic gestures), and that this applies at multiple levels such as 2, 4, 8 and 16-bar units.[18] Todd's work was based on laboratory performances, and his articles convey the impression that he is talking about a general psychological principle of expressive performance. Musicologists, however, tend to view such general principles with suspicion: perhaps the most striking lesson to be drawn from the recorded legacy of the last hundred years is the extent to which performance practices have changed. Accordingly, in a project based on recordings of

Chopin's Mazurka Op. 63 No. 3 and Sapp's visualisation techniques, I attempted to develop a way of modelling the practice of phrase arching that would make it possible to trace this historical development.

This work is based on two customised visualisations, as shown in Figures 9.10 and 9.11. The first is what I call an 'arch combiscape' and is related to Sapp's multicorrelational plots: here, however, the correlation is between the tempo or dynamic data and the shape of a rising or falling arch. The plots in Figure 9.10 consist of two triangles, the upper one representing tempo and the lower one dynamics, with time on the horizontal axis: the light flame-like patches show matches with rising arches, and the dark patches matches with falling arches. A complete arch profile is accordingly marked by the conjunction of a light and a dark patch, with the height of the patches giving an indication of the length of the arch (4 bars, 8 bars, etc.).[19] Figure 9.10a, then, is saying that Neuhaus's recording from 1955 is a perfect example of phrase arching as defined by Todd: the bilateral symmetry reveals the extent to which tempo and dynamics are coordinated with one another, as well as with the composed phrasing. By contrast Friedman's recording of 1923 (Figure 9.10b) shows very little evidence of phrase arching at all; there is a major caesura around bar 25, but a caesura is not the same thing as a regular pattern of phrase arching. In a nutshell, the story that emerges is that the elements of phrase arching exist in recordings from before the 1939–45 war, but that fully coordinated phrase arching – with tempo, dynamics, and the composed phrasing all locked together – emerges only after the war. Of course, this story is just based on an incomplete (though substantial) collection of the recordings of Op. 63 No. 3, and without doing the work one can't know how far it might apply to other mazurkas, let alone other repertories. The pivotal role of the war, however, has been remarked by scholars working on other repertories: why this might be the case is not clear, but one should probably understand the practice of phrase arching in relation not only to other aspects of performance practice at that time, but also to contemporary developments in other cultural spheres, such as architecture and design.[20]

In Figure 9.10 phrase arching at the 8-bar level is so strong that it is hard to see it operating at other levels. Figure 9.11 shows an alternative visualisation, based on the same arch-matching principle but now implemented as an Excel spreadsheet. Here the extent of tempo and dynamic phrasing at each level (2, 4, 8 and 16 bars) is shown separately. Whereas in the Neuhaus recording both tempo and dynamic arching are much stronger at the 8-bar than at other levels, Friedman's tempo arching is generally weaker and more dispersed between levels. But the values in the bar chart account for only part of the effect of expressive phrase arching. The effect of

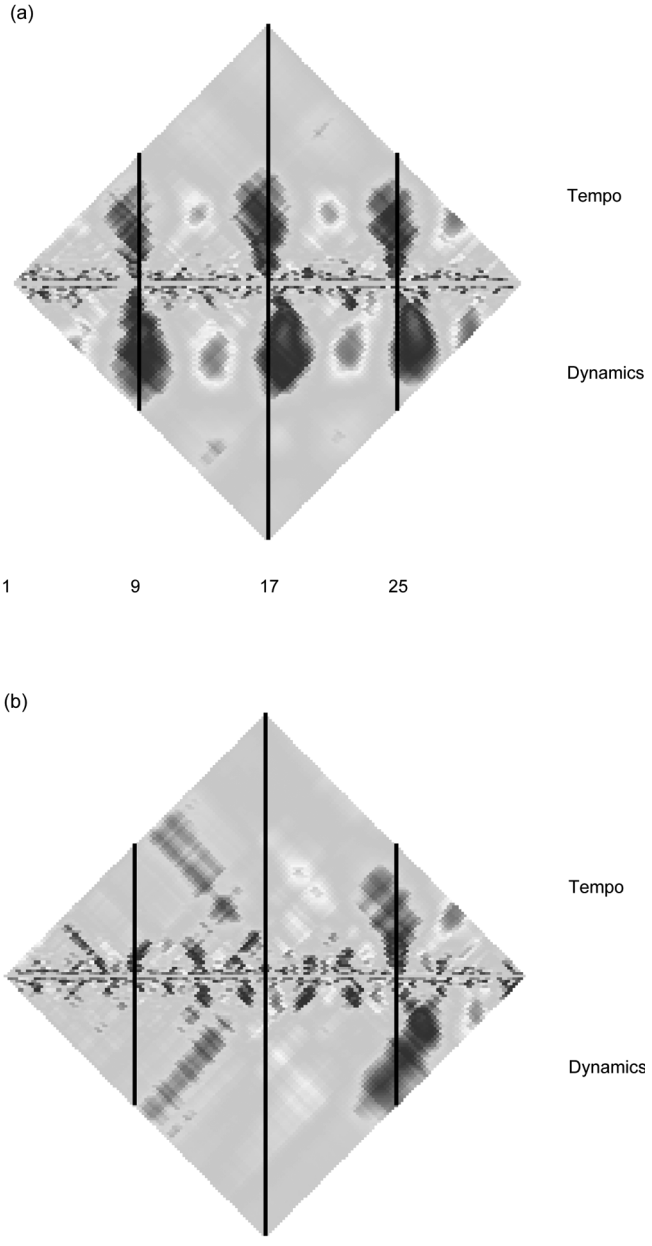


Figure 9.10 Arch combiscapes of (a) Neuhaus's and (b) Friedman's recordings of Chopin, Mazurka Op. 63 No. 3, bars 1–32: light patches correspond to rising, and dark patches to falling, arches

8-bar phrasing is so strong in Neuhaus's performance because it is reinforced by the high degree of correlation between tempo and dynamic arching, as shown in the line graph: the scale relates to this graph, with 0 signifying no correlation, and 1 signifying identity. And by combining these two distinct factors into a single formula,[21] it is possible to devise

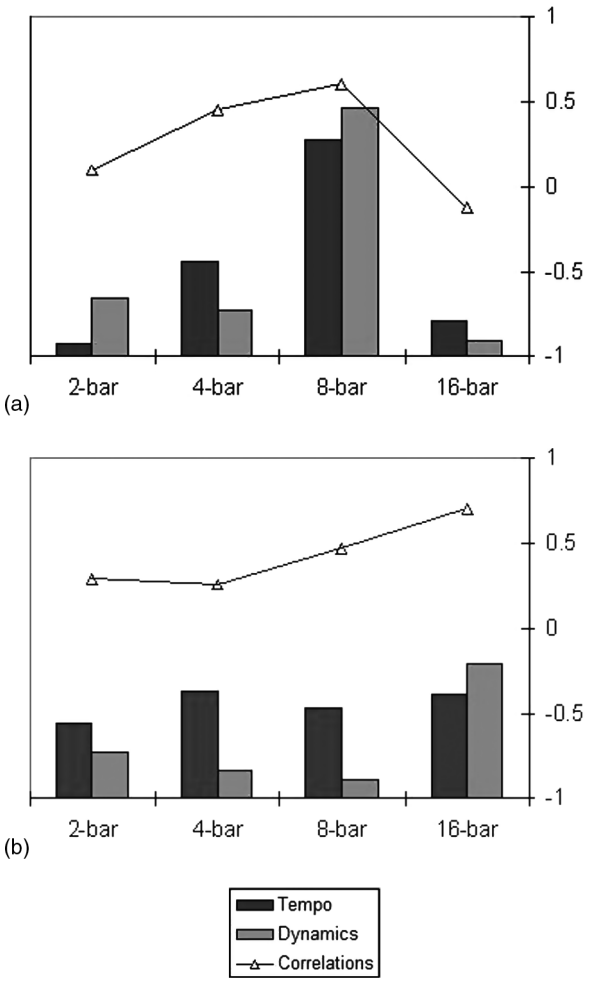


Figure 9.11 Phrase arching profiles for (a) Neuhaus's and (b) Friedman's recordings of Chopin, Mazurka Op. 63 No. 3, bars 1–32, with strength of arching and the degree of correlation between tempo and dynamic arching shown separately at each level (2, 4, 8, and 16 bars)

a rough overall measure of phrase arching. Figure 9.12 is a scattergram based on this measure, with the strength of phrase arching (highest at the top of the chart) plotted against the date of recording. One interesting finding is that the three performers of whom we have multiple recordings – Friedman, Rubinstein and Uninsky – all come out with rather consistent overall values, even though the individual profiles of their recordings vary. (In other words they achieve similar levels of phrase arching in their different recordings – rather as Grunin found similar levels of flexibility in Furtwängler's recordings of the 'Eroica' – but they do so in different ways.) Another finding is the extent to which phrase arching is associated with Russian or Russian-trained pianists (represented by the squares in

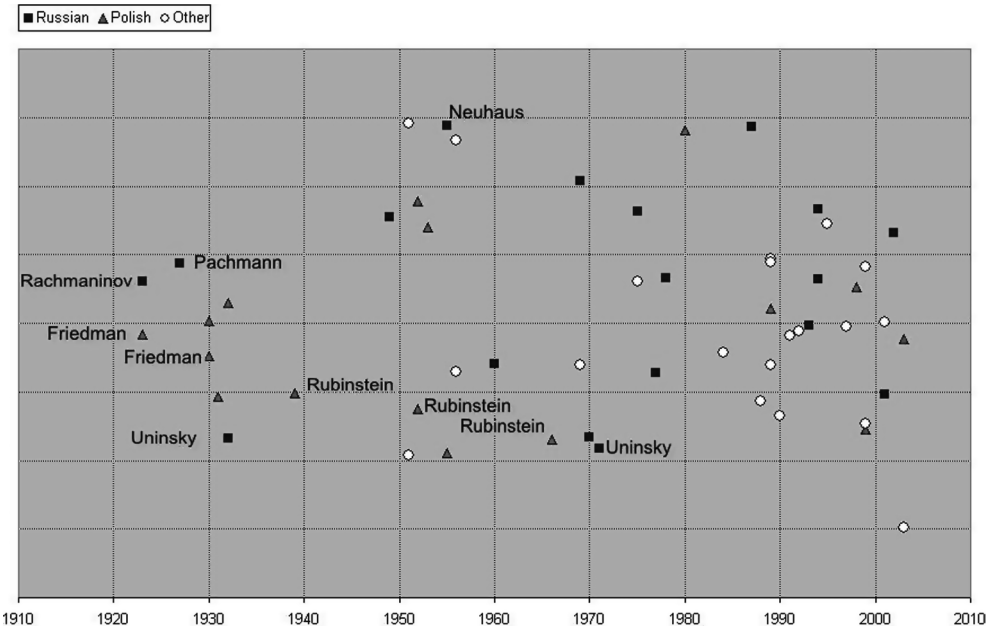


Figure 9.12 A comparative measure of the overall strength of phrase arching in recordings of Chopin's Mazurka Op. 63 No. 3, bars 1–32, plotted against date of recording

Figure 9.12): 68 per cent of them fall into the top half of the chart. Finally, the scattergram shows that, while highly coordinated phrase arching emerged after the war, performances that did not feature it continued. There is little evidence here of the narrowing range of stylistic options which many commentators have put down to the baleful influence of recordings.

The hope, then, is that analysis focused on specific features of performance will give rise to more meaningful interpretations of changes in performance style than analysis based on such undifferentiated data as total duration or overall tempo profile. It is possible to imagine a set of style-analytical tools, of which phrase arching might be one, that could be used together to characterise the style of individual performers, so facilitating the same kind of aesthetic and interpretive study of performers that traditional musicology has lavished on composers. The result would be a musicology that does better justice to music as a performing art.

The ear in culture

But can empirical, computationally based approaches such as I have been discussing really help us understand music as a cultural practice? Powerful voices have been raised both within and beyond musicology against

formalised analytical approaches in general. Richard Taruskin writes that 'turning ideas into objects, and putting objects in place of people, is the essential modernist fallacy – the fallacy of reification, as it is called. It fosters the further fallacy of forgetting that performances, even canned performances, are not things but acts.'[22] And Carolyn Abbate goes further. She claims that the experience of live performance is the only authentic musical reality and hence the only valid subject for musicology: scores and recordings – what she calls 'the tactile monuments in music's necropolis'[23] – are no more than cyphers of that experience, employed by musicologists to distance and domesticate an experience that is uncanny, unruly and ultimately irreconcilable with scholarly discourse. From such a point of view, the empirical and computational analyses discussed in this chapter must look like the ultimate sell-out.

Taruskin's reference to 'canned performances' echoes a tradition of disparaging mechanically reproduced music (which is to say music as most people today experience it) that goes back almost as far as the technology itself. But the criticisms deserve a more considered response. Analyses, whether we are talking about Schenkerian voice-leading or combiscapes, may be things, but they are meaningful only in so far as they prompt acts of informed listening: Taruskin's critique of analysis itself puts objects in place of people, the people in this case being analysts. As for Abbate, the idea of a musicology without representation is a dead end, for musical cultures are, as much as anything, cultures of representation. The way to avoid the dead hand of rationalisation – to keep music live, as the Musicians' Union slogan has it – is to understand scores and recordings semiotically, that is to say, as possessing meaning not because they are things but because they reference acts. On the one hand, recordings can be understood as the traces of performative events, whether located in a concert hall, a studio or control room (the performers in question include producers and engineers), or a teenager's bedroom. On the other hand, they are prompts to performative acts by listeners, whether in the social circumstances of a pre-war gramophone club, the domestic space of a 1950s home, a 1980s cityscape musicalised by the Walkman, or twenty-first-century clubbing culture. As we shall see, however, the distinction between trace and prompt is usually more blurred than this suggests.

The most obvious, not to say naive, way to think of a recording is as an aural snapshot.[24] On 3 April 1902, Fred Gaisberg – the Gramophone Company's first sound engineer and talent scout – visited the Vatican and made a recording of Alessandro Moreschi, who sang in the Sistine Chapel Choir and was possibly the last of the professional castrato singers. The session began with the 'Crucifixus' from Rossini's *Petite messe solennelle*, and, though the resulting recording sounds very uncontrolled to modern ears (perhaps as a result of nerves and the unfamiliar

circumstances), it was issued in the same year on the Red G & T's label.[25] This is as close as a recording can be to the trace of a performative event, although how far that event can be retrieved at this distance of time is doubtful: as the liner notes of the Pearl Opal reissue say, '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'. (Early disc recordings varied a great deal from the nominal 78 rpm.) The nearest contemporary equivalent might be the CD which the audience members of John Eliot Gardiner's Cadogan Hall (London) concert walked away with on 9 February 2006: consisting of Mozart's Symphonies Nos. 39 and 41, the CD was recorded during the first half of the concert, with 1,000 copies being made during the interval and second half.[26] Here the recording functions as a souvenir, the trace of a personal experience.

But this model obviously does not apply to recordings produced by studio multitracking, such as Mike Oldfield's *Tubular Bells* (1973), or Queen's 'Bohemian Rhapsody' (1975), the album version of which was built up layer by layer, giving it a tightness that Queen could never achieve in live performance. Nor does it apply to the Wagner recordings produced during the 1950s and 1960s by John Culshaw, who manipulated the virtual space of the stereo recording in order to compensate for the loss of the visual dimension of live opera, or Glenn Gould's 1976–7 recordings of piano pieces by Sibelius: Gould employed separate banks of microphones placed in and around the piano, alternating between and mixing the separate inputs in a manner coordinated with the musical structure.[27] Such production techniques cannot be compared to a snapshot, and indeed Gould likened his approach to film editing. Nobody who sees a film thinks it was made by leaving the camera running for two hours: films consist of the traces of a large number of performative events taking place over a period of weeks or months, edited, crosscut, and nowadays digitally manipulated, and the same is to a greater or lesser degree true of virtually all sound recordings made since tape editing became widespread. But the film still references an event or series of events of which it presents itself as a trace: it is just that the diegesis, as film theorists call it, is fictive, and understood as such by audiences. The concept of diegesis applies just as well to sound recordings, and demonstrates the sense in which the relationship between the recording (the thing, as Taruskin would have it) and the experience is a semiotic one. In other words, it is through the act of listening which the recording prompts that we understand it as the trace of an event.

And that takes me back to where I started, to the different methods of analysing recordings and purposes for doing so. In this chapter I have frequently used the words 'performance' and 'recording' as if they were more or less interchangeable, and that is because most of the time

musicologists are interested in recordings as documents of performance. But even here there is a distinction to be made. It is possible to focus quite specifically on the actual performance events of which the recording is a trace: you might do this if your aim was to reconstruct live performance practice (a difficult project, apart perhaps from the period between the introduction of electrical recording and the adoption of tape editing[28]), or in order to reconstruct the studio production processes that were involved in making the recording. But musicologists are more often interested in recordings for the listening experience they afford, and the conceptions of how music might go that they embody. From this point of view, it is not only probably undecidable but also not to the point whether a given effect was created by the performers, the producer, or the postproduction engineer: as inherently collaborative products, recordings 'are what they are', in Peter Johnson's words,[29] and it is as what they are that they circulate and are consumed as integral elements of contemporary musical culture.

One way to express this is that recordings do not so much reproduce musical performances as redefine what performance is. And if we see recordings as an integral part of a more generously conceived practice of performance, then it makes sense to apply the approaches developed by interdisciplinary performance theorists directly to recordings. Philip Auslander has written that 'to think of music as performance is to foreground performers and their concrete relationships to audiences, *rather than* the question of the relationship between musical works and performances' (this is one of the critiques of analysis from outside musicology to which I referred).[30] In saying this, Auslander primarily has in mind the extent to which performers construct – perform – identity, not generally their own identity as an individual but rather a fictive identity as an artist, what Auslander calls a persona. His point is obvious when applied to Bob Dylan or Madonna, but hardly less applicable to Karajan or Gould. And recordings play as crucial a role as live performance in such identity construction: if, in Baz Kershaw's words, it is 'a fundamental tenet of performance theory… that no item in the environment of performance can be discounted as irrelevant to its impact',[31] then this, too, applies to recordings. One obvious example concerns what Serge Lacasse calls 'phonographic staging', the creation of particular sound images through production effects such as reverberation, compression and phase shifting, as well as the stereo positioning with which Dockwray and Moore are concerned. Lacasse proposes a taxonomy of such effects, and explains that, 'rather than describing the ways in which different sound effects are *produced* in the studio, the model aims to account for these effects mostly from the point of view of the listener: how do these effects alter the

ways in which we *perceive* recorded sound sources?'.[32] And if such effects contribute crucially to phonographic performances of personae, and more generally to what recordings mean, then so do such non-auditory dimensions of recordings as cover images and liner texts, not to mention the physical and social circumstances within which recordings are experienced. No analysis of the cultural meaning of recordings can be regarded as really complete without consideration of all these matters.

Predictably, where I disagree with Auslander is over his use of the words 'rather than'. I would have preferred 'as well as'. To be sure, as I said before, an approach based exclusively on the fully fledged 'work of music', as constructed by aestheticians of musical autonomy, will be very limited in its application; actually, it will be deficient even as applied to the performance of the Beethoven symphonies, the touchstone of what might be termed 'opus composition'. But there is no either/or here. Any performance, live or recorded, can be the performance of a musical work (arguably must be for, as Bruno Nettl says, it is something close to a musical universal that 'one does not simply "sing", but one sings something'),[33] and at the same time a performance of individual or communal identity, an embodiment of the acts of actual people in concrete situations and in real time. It follows from this that there is no one way of analysing recordings, and that we should be prepared to work with as many different analytical methods as there are dimensions within which recordings signify (and one can always think of one more dimension that might be significant). But I would add that we should expect the most fruitful results when we link different, even apparently opposed, methods. The kind of computer-assisted close listening I described in the first part of this chapter can refine an analysis of identity construction in terms of phonographic staging as well as of the expressive effects through which performers create meaning; conversely, consideration of the performative effects and social consumption of recordings provides a context within which to make sense of observations resulting from close listening or computational evaluation. In short, cultural analysis can be supported by empirical analysis, and empirical analysis given purpose by cultural analysis. It's a win–win relationship.