

Expressive Gestures in Schubert Singing on Record

Daniel Leech-Wilkinson

1

Musicology and performance

Musicology, as is well known, has tended to think of music as existing in works, and works as existing in scores: scores thus provide the hard evidence from which music may be understood. Philosophy has done much to tighten up those beliefs, giving them a more rigorous intellectual basis.¹ But times are changing, and a new generation of musicologists finds this view of music far too narrow.² Increasingly, too, there is a recognition that, while musicology has been aiming at objectivity, its findings are in fact shaped by musicologists' personalities and their backgrounds as much as by any hard evidence.³ There's also an increasing understanding of the ways in which popular music and world musics challenge musicology's definitions of music.⁴

What I think has not been widely noticed yet, although it's going to become more and more obvious over the next few years, is the extent to which hard evidence from 100 years of

¹ To mention only some recent relevant examples, Peter Kivy, *Sound and Semblance* (Princeton University Press, 1984); Stephen Davies, *Musical Meaning and Expression* (Ithaca: Cornell University Press, 1994); Nelson Goodman, *Languages of Art* (Hackett, 1968); Jerrold Levinson, *Music, Art, and Metaphysics* (Ithaca: Cornell University Press, 1990); and *Music in the Moment* (Ithaca: Cornell University Press, 1997); Stan Godlovich, *Musical Performance: a philosophical study* (Routledge, 1998). For a view with which the points made here can be more easily reconciled see Aaron Ridley, *The Philosophy of Music: theme and variations* (Edinburgh University Press, 2004).

² For a convenient sampling of recent thinking see ed. Nicholas Cook and Mark Everist, *Rethinking Music* (Oxford University Press, 1999).

³ See for example Kevin Korsyn, *Decentering Music: a critique of contemporary musical research* (Oxford University Press, 2003). For specific case studies see Daniel Leech-Wilkinson, *The Modern Invention of Medieval Music: scholarship, ideology, performance* (Cambridge University Press, 2002).

⁴ As a starting-point see Cook and Everist, *op. cit.*, *passim*; and ed. Georgina Born and David Hesmondhalgh, *Western Music and its Others: difference, representation, and appropriation in music* (Berkeley, University of California Press, 2000).

recorded musical performance also undermines many of the old certainties about music's identity.⁵ When one listens closely to, and thinks about, the huge diversity of manners of performance over the past century, it's impossible to avoid the conclusion that pieces of music change over time. The conventional musicological response would be to lament this, to regard this gradual change as an accumulation of 'errors' that should be 'corrected' by returning to the original sounds and texts.⁶ One can argue for ever about this in principle, but in practice it's a nonsense. We don't have the evidence from before the age of recording, and in any case, performance styles reflect the mood of their time, and begin to sound quite peculiar when that mood has changed.

I can't provide sound examples with this article, but I invite readers to listen to an early twentieth-century recording of just about any singer of opera or Lieder. Lilli Lehmann singing Schubert's *Du bist die Ruh*, would be a fair example, or Elena Gerhardt singing *Die Forelle*.⁷ These performances are fascinating to hear and to study; but we can't sing like that any more. And I really mean that. I don't believe it can be done convincingly, however hard you try. Brave and impressive results have been achieved by instrumentalists imitating early twentieth-century performance styles,⁸ but they never sound quite right, and of course for singers even imitation seems out of the question because so much that is essential to voice production would have to be changed. But my point is that, even if these technical obstacles could be overcome, one cannot make music convincingly with somebody else's performance style because performance style, as we shall see later in this article, translates into sound manners of emotional communication that change over time. To adopt another period's performance style would therefore involve relearning how to communicate one's feelings,

⁵ To date the best discussions of early recorded performance styles are Robert Philip, *Early Recordings and Musical Style* (Cambridge University Press, 1992); his *Performing Music in the Age of Recording* (New Haven: Yale University Press, 2004); and Timothy Day, *A Century of Recorded Music: listening to musical history* (New Haven: Yale University Press, 2000)

⁶ For a recent, though to be fair a much more nuanced statement along these lines, see Peter Walls, *History, Imagination and the Performance of Music* (Woodbridge: Boydell Press, 2003).

⁷ At the time of writing Lilli Lehmann's 'Du bist die Ruh', originally issued on Odéon 50432 (matrix XB 3137), is available in CD transfer as 'Lilli Lehmann (1848-1929)' Preiser, *Lebendige Vergangenheit*, MONO 89185, track 20. The Gerhardt performance, originally issued on HMV DA 835 (matrix Bb9436-1) is available for free download from <http://www.charm.kcl.ac.uk/> and <http://www.kcl.ac.uk/music/ksa/> .

and to learn to do it in a language that one's contemporaries could not easily understand. In other words, it would be to make incomprehensible something that it is essential to our happiness to make as comprehensible as possible. The instinctive barriers to that are, I suggest, too strong to allow us to do it in a thoroughgoing way. We shall always wish to hang on to vestiges of our own period's style, not just because of our fear of being unable to communicate with others but, more essentially, because we would otherwise be unable to communicate with ourselves. How could we know that we had arrived at a coherent and persuasive manner of performance if we couldn't feel it as coherent and persuasive in ourselves? Music communicates feeling, and if it doesn't it doesn't sound like music. Our musical capabilities are thus rooted in the social style of our time, and most particularly, in the style in which we conventionally express our feelings. There will be more to say about this below. But even now we can say with absolute certainty, after listening to these early twentieth-century performances, that Schubert's *Du bist die Ruh* and *Die Forelle* were not the same pieces then that they are now. The notes may have been the same, but they felt quite different.

Recordings show us that musicology has forced us to think of music as something much more concrete and fixed than it is. Promoting this distorted view of music has been hugely to musicology's advantage. By constructing the imaginary musical 'work', and limiting the notion of music to just that, it becomes far easier to talk about music in what sounds like a disciplined and comprehensible way.⁹ By making music a text instead of sounds one has something to look at, something to measure, something to 'see' patterns in, something to reproduce exactly, and also something to sell, to copyright, to own. Above all, one has something that can be precisely fixed.

But one does not have music. Music is none of these things. You can't see it, you need a machine to measure it, you need a machine to store it in order to examine it for patterns and in order to reproduce it, to own it, copyright it and sell it. The human body can't do these things on its own. And this, more than anything, is why music has been so hard to understand. It's true that we've had machines that will store it, and storage media that we can own, copyright and sell, for 100 years. We've had some of the technical means to measure it

⁸ Hear the CD accompanying David Milsom, *Theory and Practice in Late Nineteenth-Century Violin Performance* (Aldershot: Ashgate, 2003).

⁹ On the construction of the work concept see Lydia Goehr, *The Imaginary Museum of Musical Works: an essay in the philosophy of music* (Oxford University Press, 1992); and her *The Quest for Voice: music, politics, and the limits of philosophy* (Berkeley: University of California Press, 1998).

for most of that time. But it's not been easy. And given the huge weight of convenient cultural assumption about music's identity with the imaginary work, it's not really surprising that it is only now, with the universal availability of cheap computing, that it's become possible for a new generation of musical scholars to study music as sound, as what it is. Now that enough people have the technology on their desks there is a real chance that the academic view of music might at last change to reflect more accurately the view of music that non-academics have always had, namely that music is what happens when you hear it.

I think this is an immensely important development. The split between musical academics and everybody else who deals with music has been widening for 200 years precisely because the professionals have increasingly adopted a view of music that is professionally fruitful but radically at odds with the view of performers and listeners. Now we have ways of talking about the sound of music, that gap can be closed. It's always good to remember that there are millions of people who have powerful musical experiences without knowing any of the things that musicology knows. It seems to follow that if, as musicologists, we were hoping to enhance people's enjoyment of music then we've been looking at the wrong things. Music happens in the interaction between sounds and the brain. If we want to understand music that's where we have to look for it. So far, this has been the province of the sciences: psychoacoustics, psychology, neuroscience in particular, and in the last few decades enormous progress has been made. Musicology has contributed almost nothing, except in so far as – and this may be to a larger extent than we realise – experiments have been set up and analysed under the influence of things that musicology has taught about how music works. I'm thinking especially of experiments that relate listener or performer response to musical structure. Are we sure that musicologists' emphasis on musical structure, and their understanding of it, is correct? Does it really reflect how music is perceived, or are we in a vicious circle in which musicology tells science what kinds of structures exist, and science sets out to find them reflected in peoples' responses? We have to be a little bit careful here. It would seem perverse to doubt that, for example, cadences are matched with various performance strategies for pointing them up in sound, strategies to which listeners respond. There is overwhelming evidence now that this is so. One senses how notes work within phrases, one shapes those notes in performance so as to emphasise the phrase structure, and listeners respond to that. Experimental work with performances is absolutely right to pay attention to the way music works at this level. Musicology's beliefs about how melody and harmony work at this level – the level of which we are most fully aware as we listen – are surely well founded. But it's far from certain that musicology has been correct to extend those principles of organisation over much larger spans of time. It's a nice idea that the same principles as operate in a phrase operate over a whole piece. But is it true? For composers it may be, especially since composers began to be trained by musicologists. But is it true for

listeners? The few empirical studies cited by Gabrielsson and Lindström suggest not.¹⁰ So we need to keep minds as open as possible about which musicological beliefs are really grounded in perceptual fact.¹¹ A good case could be made for an extensive research program to test beliefs about large-scale harmonic and formal structures. Do we perceive them or don't we? Does it matter whether there is a recapitulation, or whether the piece ends in the key in which it began, and so on? To borrow Ian Cross's word, there is a huge amount of 'folklore' about music that is considered absolutely basic by composers and academics and that is taught to every child who studies music, and believed by them throughout their lives, that may be quite wrong. Identifying it is urgent, and it's one of the things we can do together. I put it like that because what must be overwhelmingly obvious from all this, is that the kind of musicology I envisage, and for which I think the time has now come, is a musicology that is to all intents and purposes merged with studies of music perception.

2

Recordings as evidence

If musicologists want to study music as sound we have little choice, it seems to me, but to work with recordings. Recordings store performances and allow us to examine moments within them, to hear them repeatedly, to think about them, and to analyse them in detail. But if recordings are to provide our data then we're going to need to be reasonably confident, first of all, that they represent performances sufficiently accurately to do duty for them. As with any source of evidence, their production and original function need to be understood before we can understand what they transmit: how were these sources made and why? Above all, we need to know what a recording is of. And to do that we're going to have to learn rather a lot about the history of recording, its economic, social and technological history, all of which have a significant bearing on what comes off a record when we play it.

¹⁰ Alf Gabrielsson and Erik Lindström, 'The influence of musical structure on emotional expression', in ed. Patrik N. Juslin and John A. Sloboda, *Music and Emotion: theory and research* (Oxford University Press, 2001), 223-48 at 234.

¹¹ For a very welcome antidote see ed. Andrew Dell'Antonio, *Beyond Structural Listening? Postmodern Modes of Hearing* (Berkeley: University of California Press, 2004).

This is not the place to go through all that in depth.¹² But I will summarise some conclusions that follow, I think, from a detailed consideration of the changing relationship between performance and recording over the past 100 years.

Cylinders and 78rpm (shellac) discs (acoustic era, 1877–1925)

- do preserve single unedited takes
- can transmit accurate relative timings and relative pitches (subject to speed change during disc cutting)
- cannot transmit precise data on speeds or pitch
- cannot be used as evidence for the colour of a voice, still less an instrument, because they don't transmit the higher frequencies that determine it, and the recording equipment tends to alter the balance of those frequencies that get through to the disc
- do not represent concert-practice scorings in instrumental music, nor concert balances between instruments, nor venue acoustics

Cylinders and 78rpm discs (electrical era, 1925–ca. 1948) improve on acoustic era recordings in these respects:

- can transmit more information about colour, though not much more
- do represent concert scorings, and are more likely to represent concert balances, and are much better at representing recording venue acoustics

LPs (vinyl) and analogue tape (ca. 1948–ca. 1970)

- do provide a relatively faithful representation of sound colour
- do provide reliable data on speeds and pitch
- because of tape editing cannot be relied upon to represent one continuous performance
- are unlikely to represent the balance of instruments/voices heard by a listener at the recording session

¹² A fair working knowledge of the changing history of recording can be had from Roland Gelatt, *The Fabulous Phonograph* (London: Cassell, 1956) and (more briefly) Peter Copeland, *Sound Recordings* (London: British Library, 1991); supplemented with detail on the early period provided in George Brock-Nannestad: 'The Objective Basis for the Production of High Quality Transfers from Pre-1925 Sound Recordings', Preprint of paper delivered to the Audio engineering Society's 103rd convention, 1997 (AES Preprint 4610); and his 'Authenticity in the Reconstruction of Historical Disc Recording Sessions', Preprint of paper delivered to the Audio Engineering Society's 105th convention, 1998 (AES Preprint 4829).

Digitally recorded LPs, CDs and subsequent media

- may or may not provide a faithful representation of sound colour
- may or may not provide unaltered data on speeds and pitch
- are unlikely to represent one continuous performance for more than a few seconds at a time
- may or may not represent the balance of instruments/voices heard by a listener at the recording session

General notes

‘Live recordings’, whether on LP or later, may have been assembled from recordings of several continuous performances

LP or CD transfers from 78s almost never represent faithfully the content of the original discs.

Now you may feel, after a list like that, that using recordings as evidence for performance is absolutely hopeless. Certainly we need to be very careful. There is no point in trying to draw conclusions about dynamic shadings in a performance recorded on shellac. There is no point in talking about large-scale tempo control in a performance of, for example, a Beethoven piano sonata transferred from shellac to CD, unless you’ve consulted the original shellacs and know where the side-breaks were and how they were managed. And so on. One needs to know these things. But – and it’s a very big but – even though recordings do not literally represent performances they do everything that performances do. Which is why records of classical music sell, and why so many of us have shelves groaning under the weight of LPs and CDs. Recordings sound like performances, and since sounding is what performances do, to sound like a performance is to be a performance. Let me just unpick this statement so that there’s no misunderstanding. I’ve just listed many ways in which recordings are not recordings of performances. That’s true if we define a performance as an occasion, a social situation in which people assemble to perform and listen to music. But if we define a performance as the persuasive sounding of a piece of music then the value of recordings looks quite different. Is this an adequate definition? For the listener I think it is. (And notice that I say the listener, not the spectator. There are advantages but also disadvantages to being present at a performance and so to having visual input also available.) From the listener’s point of view a performance of a piece of music is the sound of that piece from beginning to end. And it’s the listener’s point of view that I want us to be able to understand. Because music is for listeners and no one else. Listeners include the performers, obviously, who may indeed be the only listeners if there is no audience, but performers certainly listen, and more closely than anyone. And listeners include score readers, who *hear* a performance in their heads, a performance made by themselves.

Let's look at this from another angle. The crux of the problem with recordings is often said to be in the editing. A producer assembles a performance at the editing stage from fragments of many separate takes. In that it's manufactured other than by the performer in real time, the result looks like an artificial construction. But it doesn't sound like one. The editor is a musician, often a very sensitive one, and assembles a performance that sounds persuasive, even more persuasive, in fact, than any that the performer gave in the studio. How is this possible? Quite simply because of the unfixedness of music. Wonderful musical performances can be made via editing precisely because it is the nature of music – part of what it is – that there is no ideal way in which one moment follows another. At any moment a performer has choices to make, but choices from possibilities that are not fixed in advance. They are limited but not predetermined. They are limited to some extent by perceptual rules, and to a greater extent by current taste. (I'll provide examples of both these later.) But even within those limitations (whose limits, incidentally, are currently impossible to determine), even within those limitations a performer can decide, largely instinctively and on the spur of the moment, what to do next. Consequently, a performance edited together from numerous different choices need be no less convincing than a performance given as a continuity. If it sounds convincing it is a performance, one that does what live performances do from moment to moment.

Seen from this, perceptual, point of view, most of the objections one might have to using recordings as evidence for performances fall away. One still needs to be aware of how they were made, but only in order to make a distinction between statements about what the performers did and statements about the musical result that we perceive. Provided that the performers might have done the things that we hear from the record, even that distinction loses much of its importance.

3

Expressive performance

I've emphasised already that music is to a surprising extent unfixed. It can be different in many ways in every performance and yet be musical. There are limits, however. A useful demonstration comes from discs with audibly bad edits. They are not common, of course, though curiously they are more common on modern CDs than they were on LPs, despite the fact that (good) digital editing can be impossible to spot, perhaps because editors don't treat computer files with the same reverence as the single master tape that was previously the end product of their work. Whatever the reason, poor edits are valuable, because they offer

evidence of the limits of acceptable flexibility in performance. An example can be found on the disc from which I am going to draw most of my examples below. This is a Deutsche Grammophon CD of Kathleen Battle accompanied by James Levine in songs by Schubert, and readers who are able to acquire a copy will find it helpful to have on their CD players as they read.¹³ This DG disc is very heavily edited. On the minus side that means one has to use it with great caution, but on the plus side it means we can learn something of what's gone on. Battle's and Levine's performance of *Nähe des Geliebten* includes an edit that doesn't work because it's not plausible and wouldn't happen in any live performance currently imaginable.¹⁴ (I say 'currently' because the history of changing performance on record suggests that we be wary of assuming anything is fixed until it's shown to be perceptually necessary.) Plausibility is the crucial test. If we can imagine something happening live then we can accept it as genuine, whether or not we like it. Music Example 1 shows the notation of the relevant passage, annotated with the length of each quaver. (The offending edit comes right at the end.)

[Example 1 near here]

What's happened is that the final quaver of bar 5 (1' 09" on the disc), setting 'In' on the high G, has lost 0.2–0.25 seconds from its end. The expected length can be worked out by what leads immediately up to it. In the preceding bars there's a pattern set up of longer sixth quavers: 'nen' of 'fernen', 'der', 'sich'; the end of the bar seems to be setting up the same *ritenuto*, but sooner. The quaver before 'In' is already as long as the quaver at 'sich', though it's not yet the end of the bar, so the last quaver must be expected to be longer still, probably about 0.6–0.65s, certainly not 0.36. And it sounds as wrong as it looks. The note is just too short to be plausible in that context. This example illustrates just how small a departure from expectations can disrupt the effect of a performance: no more than a quarter of a second in fact. This seems paradoxical when one considers how wide the range of possible variation is in timing and amplitude at any moment in a performance. Clearly some kinds of changes are artistic and some are disastrous, and it's not a matter of how great they are but rather of how they fit into their context.

By contrast, consider just about any performance of any piece of classical music in its treatment of a structurally significant cadence. The example I have in mind as I write is

¹³ 'Schubert Lieder', Kathleen Battle and James Levine, DG 419 237-2 (recorded 1985 & 87). At the time of writing the first version of this study it was still available, but it has since been withdrawn along with much of DG's classical list.

¹⁴ *Ibid.*, track 9.

Benno Moiseiwitsch playing the D-flat major Prelude of Chopin,¹⁵ but just about anything would do. If I use a digital editor to alter the length of the penultimate chord, making it longer in one version and shorter in another I can produce several versions with ritardandi at different rates of deceleration, none of them Moiseiwitsch's, but all of them making reasonably good sense. The difference between the lengths of the cadential notes can be considerably more than the difference between Battle's edited note and what she is likely to have sung in the studio, but all the piano examples are musically plausible, whereas her edited note is not. Moiseiwitsch is playing longer notes, and spreading the cadence over a longer period of time, so differences in note length are less obvious; and in addition we expect flexibility at cadences, often a lot, sometimes hardly any. It may be that the proportional relationship between the notes in a cadential preparation matters more than the actual lengths: however long or short there is usually a rate of deceleration that we can follow without surprise. So although these two examples, Battle and Moiseiwitsch, seem to have opposite results for us, they are both illustrating the same principle. Music establishes a context as it goes along within which certain timings, or pitches, or loudnesses are going to seem plausible – in other words are going to make 'musical sense' – and others will not. Exactly where the limits lie has yet to be shown, but it wouldn't be too hard (only very time-consuming given the range of repertoires that would need testing) for a psychologist to design an experiment that tested the limits within which various edited note durations are plausible; and one could probably do the same for amplitude, and for various kinds of pitch modifications. It would be rather useful. But one would have to bear in mind that the results might not be true for all performance styles: things that sound plausible in early twentieth-century recordings, once one has got used to them, might not sound plausible in performances from the 1970s. And of course one would want, also, to test that phenomenon of 'getting used to' different performance styles. There are many things musicology is starting to notice about performance styles on record that could be pinned down more precisely through controlled experiments; but this notion of performance style is one of the most interesting, because its essential role in the identity of music has not yet been fully appreciated.

To explain what I mean we need to go back a step. It's well understood in the music psychology literature that a literal performance of the notes in a score is not a musical performance. But it's always useful to be reminded of what a literal performance sounds like. I invite readers to download and play a literal MIDI transcription (just the pitches and

¹⁵ Originally HMV C 777? (matrix 2EA 135??) [7772-5, 2EA 13517-24, CHECK KSA]. Reissued on Appian APR 5575, track 15, and Testament SBT 1196, among others. On the Appian transfer the cadence in question is at 1' 06".

durations) of the music notation for this same Chopin Prelude,¹⁶ and, if you can, to compare it with a pianist's recorded performance. For musicians, listening to the MIDI transcription, in which every quaver is exactly the same length, is intensely uncomfortable; it almost hurts. But as soon as one hears the human performance one relaxes: it comes as a huge relief. I find those physical responses rather interesting. What is it about the irregularity of the human example that makes us comfortable? How it's produced in performance is now beginning to be understood, thanks to work by Eric Clarke, Bruno Repp and many others.¹⁷ It's well understood that music-making, musicianship, musicality (there are several words in English that refer to more or less the same phenomenon), as opposed to sounding the written notes, involves performing notes at lengths, amplitudes, or frequencies other than those indicated in the score; and it's well understood that there is some kind of grammar that determines what kinds of variations are desirable in particular circumstances within the score. Certain functions within a musical structure need to be emphasised by drawing attention to a note and away from its neighbours. Less well understood, I think, at any rate in the scientific literature, is that the implementation of this grammar changes over time. For example, a prominent note in a score that in 1910 was emphasised by sliding up to it from the note below, in 1950 might have been emphasised by vibrating on it, and in 1990 by increasing and then decreasing its amplitude.

What we're dealing with here is a really fascinating phenomenon; a large-scale shift in the way that a perceptual system in the mind is engaged in making sense of incoming data. Functionally the data remains the same – that is to say, the importance of certain notes continues to be signalled – but the way those functions are signalled changes, and over the years it changes a lot. At any one time the signals are relatively stable in form, and if one is only familiar with the current vocabulary of signals then, as I tried to show in the first part of this article, it is almost impossible to imagine performing with a different vocabulary, and it's hard to take seriously a performance that uses another – which is why people at first laugh when they hear very old recordings. But we've also found, now that examples of performances from over 100 years are easily available all together, that we can get used to

¹⁶ The example I used can be downloaded from www.kcl.ac.uk/music/dlw/dflatprel.mid .

¹⁷ Among many other studies see Eric Clarke, 'Rhythm and Timing in Music' in ed. Diana Deutsch, *Psychology of Music* (New York: Academic Press, 1999), 473-500; Bruno Repp, 'A Microcosm of Musical Expression', *Journal of the Acoustical Society of America* 104 (1998) 1085-1100, 105 (1999) 1972-88, 106 (1999) 469-78; Renee Timmers, *Freedom and Constraints in Timing and Ornamentation: investigations of music performance* (Maastricht: Shaker Publishing, 2002).

multiple systems. And, if you look at very recent performances on record, you can see that we're also capable of mixing vocabulary from different periods in a way that works, creating a new style with old ideas.

We need to understand much more about how this is possible. Quite a good way of conceptualising a vocabulary of musical signals is to invoke the notion of a performance style, composed of a collection of smaller gestures, rather like a conversational style, or a style of acting, both of which one could study through recordings from the past 100 years, and come to understand as changing languages for the representation of mood and emotion. Another way would be to use the language of semiology, of course, but I'm deliberately not doing that because it's been used extensively in musicology to talk about features of scores, and I particularly don't want what I'm talking about to get confused with that, although it is related and the relationships will need to be explored. But many of the attributes given to scores by musical semiologists are in fact attributes of performances of scores, not of the scores alone, and that's not been adequately realised.

So a performance style can be thought of as a collection of gestures in use at a particular time. What is a gesture? It's an irregularity in one or more of the principal acoustic dimensions (pitch, loudness, duration), introduced in order to give emphasis to a note or chord, usually the start of a note or chord. In length a performance gesture can extend over several notes, but usually applies to one, often only to part of a note, and may be as short as a few tenths of a second. A gesture is shaped by details of a note that are not specified in the score. It's identified by its difference from its surroundings; and by its giving the note which it shapes the expression of some function or meaning. Ultimately it is that last point that is the most important. Performance gestures make musical scores expressive of something. How they do it, and what they are expressive of are key research questions, because making notes expressive of something is making music – not just organised sounds, but organised sounds that communicate. Since music has no semantic meaning, understanding the process of communication is the key to understanding music.

So let's look at some examples of expressive gestures. My first comes from Kathleen Battle's singing of the beginning of Schubert's *Die männer sind méchant* (Music Example 2)

[Example 2 near here [notation]]

In the text, by Johann Seidl, a girl tells her mother that she was quite right to distrust her daughter's lover; she has just seen him kissing someone else. She begins, "You told me so, Mother, 'He is tearaway'" ('Er ist ein Springinsfelt'.) Knowing the text, it is fairly easy for

most listeners to agree that this passage makes a general impression of disgust, an emotion that Gabrielsson and Juslin tell us has rarely appeared in empirical studies, perhaps because it needs more clues than just sound to identify it unambiguously.¹⁸ Here the text and sound together leave little room for doubt. But what produces that effect? One could describe it in general terms, but in fact it's possible to say confidently what is going on in the sound here because one can see it clearly in a spectrum analysis. For this I've used a software program for PC called Spectrogram, but many others would do just as well.¹⁹ The display shows all frequencies above a set dB level, arranged on the X axis, plotted against timing on the Y axis, with loudness shown by colour. So it's possible to see and measure accurately every sound that is present in a performance: it's a convenient aid to precise listening. In Battle's treatment of 'Er ist ein Springinsfeld', 'Er' starts loud and slightly sharp (suggesting a shrill manner of speech). 'Er' is connected by a portamento slur to 'ist', so 'Er' seems hissed by association ('Errissst'). 'Ist' is dead straight, the pitch sounds for less than a third of length of the beat, the rest is 'sss'. The sharp 't' of 'ist' is attached to 'ein' and gives 'ein' some bite, which it couldn't otherwise have; and Battle also gives it a nasal tone and a very fast upper mordent ('e/in'). 'Sh' (of 'Springinsfelt') is all higher frequency and starts early, exploding into a loud 'pring' which starts straight, then slides in a continuous portamento down through 'ins' to 'feld'. And there's a hairpin crescendo-diminuendo through 'insfeld', giving a thrown-away end to the word that perhaps suggests dismissal.

These are all expressive gestures, evoking sounds from life. How do they work? Some are onomatopaeic, the 'issst' which we've learned to associate with hate and in particular, through its associations with geese and snakes, a threatening hate that could lead at any moment to violence. Then there are the sudden explosive consonants, 'T-ein' and 'shPRing', evoking the sound of sudden violence. And the rapid diminuendo at the end of a phrase that evokes a dismissive turning away. They're signs of emotive actions we recognise from daily life. This isn't a separate musical sign-language, in other words, but rather uses the mind's naturally selected ability to connect phenomena through common features. It's a survival skill that's become a way of understanding the world. The performer integrates the music with our emotional lives. And that's surely one of the most important things that performers do, and one of the absolutely essential ways in which music works. Of course singing

¹⁸ Alf Gabrielsson and Patrik N. Juslin, 'Emotional Expression in Music', in ed. Richard J. Davidson, Klaus R. Scherer and H.Hill Goldsmith, *Handbook of Affective Sciences*, (Oxford University Press, 2003), 503-34 at 528.

¹⁹ Available for download from www.visualizationsoftware.com/gram.html .

teachers have been talking in these terms for ever, but musicology has so far shown no interest in learning from them.

Many kinds of gestures in song come direct from speech. The meaning of a sound gesture in speech is transferred to singing by converting the speech gesture into something with a more precise pitch and duration determined by the requirements of the musical context. Often the speech gesture invoked is determined by the text, but by no means always. There's an example in Battle's recording of Schubert's *Rastlose Liebe*,²⁰ where she sings 'Ohne Rast und Ruh' (without peace or rest). In Battle's wildly restless performance 'ohne' is the word given the largest gesture in the phrase, not because it is the most important word but because Schubert's setting puts it in the most expressive place, and the singer can get the strongest effect by working with that. This emphasises that gestures borrowed from speech don't necessarily have to arise from the text in order to work; their expressive content in speech may be taken over and applied in music, either because their emotional content works effectively in a musical context or simply because the sound seems right. And for that same reason I suspect that speech-derived gestures may also be used in instrumental performance. I'm not going to take space to look at examples of that now, but it's a possibility to which I hope to return in another study.

What about the general character of Battle's voice in this performance? This is not how Battle normally sounds in Lieder. A more typical extract, which conveniently includes both her lyrical and her characteristic clipped styles would be *Lachen und Weinen*.²¹ So what's different about *Rastlose Liebe*? Spectrograms, naturally, show vibrato and allow one to measure it easily. And so one can say that in *Rastlose Liebe* Battle doesn't just make 'Ohne Rast' a continuous portamento – we can hear that – she also changes her vibrato, speeding it up by around 10% from her usual rate (which is very roughly 135ms per cycle as opposed to 150 here) and reducing its width by around 30%. The effect is that she sounds terrified. This is a neat example because it shows so clearly how effects that we all recognise immediately as signs of terror – racing heart, tremor in one's voice – can be reproduced analogously, indeed almost literally, in singing, and they inevitably cause us to share some of the feelings that would normally evoke them. There's lots of research along these lines. The motor theory of speech perception,²² Juslin's functionalist perspective,²³ Sloboda's 'dynamic awareness',²⁴

²⁰ Battle & Levine, track 14.

²¹ Battle & Levine, track 11.

²² Alvin M. Liberman and Ignatius G. Mattingly, 'The motor theory of speech perception revised', *Cognition* 21 (1985), 1-36.

Cox's mimetic hypothesis,²⁵ Watt & Ash's hypothesis that 'the action of music is to mimic a person',²⁶ and indeed, Peter Kivy's contour theory,²⁷ further developed by Stephen Davies,²⁸ are all describing this phenomenon, and it seems absolutely clear to me that this is a fundamental key to understanding musical communication. We read sounds through what our bodies would do to make them. The truth of that is particularly clear in this example, because the causes and effects are so obvious and easy to identify. But the same process is likely to be working in much less obvious cases as well.

My next few examples all come from Schubert's *Die junge Nonne*, whose psychological portrait of a disturbed young woman offers us the chance to study the vocal representation of fear in more detail. (There isn't space here to discuss what this text, by Craigher de Jachelutta, is really about, but Schubert's musical materials, especially the ever-present storm motif, clearly suggest that its profile of a girl joining a convent in order to be married to God is not to be taken at face value.)

Wie braust durch die Wipfel der heulende Sturm!	How the howling storm rages through the tree-tops!
Es klirren die Balken, es zittert das Haus!	The rafters rattle, the house quakes!
Es rollet der Donner, es leuchtet der Blitz!	Thunder rolls, lightning flashes,
Und finster die Nacht, wie das Grab!	the night is dark as the grave!

²³ Patrik N. Juslin, 'Communicating Emotion in Music Performance: a review and a theoretical framework', in ed. Juslin & Sloboda, *op. cit.*, 309-37.

²⁴ John Sloboda, 'Does music mean anything?', *Musicae Scientiae* 2 (1998), 21-32; reprinted in Sloboda, *Exploring the Musical Mind: cognition, emotion, ability, function* (Oxford University Press, 2005), 163-72.

²⁵ Arnie Cox, 'The Mimetic Hypothesis and Embodied Musical Meaning', *Musicae Scientiae* 5 (2001) 195-212.

²⁶ Roger J. Watt and Roisin L. Ash, 'A psychological investigation of meaning in music', *Musicae Scientiae* 2 (1998) 33-53.

²⁷ Peter Kivy, *Sound Sentiment: an essay on the musical emotions, including the complete text of The Corded Shell* (Philadelphia: Temple University Press, 1989).

²⁸ Stephen Davies, *Musical Meaning and Expression* (Ithaca: Cornell University Press, 1994), especially chapter 5.

Meta Seinemeyer, in a recording from 1928,²⁹ contrasts the opening phrases, in which key words (Wipfel, Balken, Donner) are hit hard through initial consonants sung at full amplitude, with the softer-edged 'Und finster die Nacht', whose sounds crescendo up to their full strength (which is less than for the hard notes). She also slows down for them, but that's less crucial. The speech analogy is obvious: spitting-out sounds evoke anger in speech and by analogy the fury of the storm; crescendoing sounds evoke something more complex, since there are a number of situations in which we might crescendo through a sound in speech. Mystery tinged with fear might be one, and is perhaps what is evoked here. But one can also think of the contrast between hitting and stroking or pushing. Thinking visually one would call on hard edges contrasted with blurred. All these are obvious equivalents using different senses, equating to the perception of this passage of sound. To pin a precise meaning onto these sounds at 'finster die Nacht' would be silly, because it would be to attempt to make precise something that is by its very nature not precise. That is the point of the gesture, that it evokes unease, which by definition cannot be precisely explained: its precise meaning is imprecision. But my methodological point is that once one understands what's going on in the sound it's not hard to see what it means.

Kathleen Battle in the same passage offers us several details that suggest fear through loss of control.³⁰ In 'zittert das Haus' Battle sings 'das' on a different note entirely from Schubert's, and not a scale note: it's 650Hz instead of 550 (c"#), about a tone and a quarter above pitch. Then at 'finster die Nacht' she slides up to 'finster' and 'Nacht' are obvious enough to the ear; a little more subtle – and this is another case where a visual display can help – is the shallow but longish slide up to 'wie' and 'Grab', and the combination of that on Grab with increasing vibrato and crescendo/diminuendo. These are not characteristic of Battle's normal style: in fact she's extraordinary among female singers of her generation for her ability to start a note with all the amplitude and vibrato it's ever going to have, and for keeping them both absolutely regular throughout a note; it's an exceptionally regular voice, and so these small changes have much more significance for her than they would in others' performances. One has to read gestures in relation to their local context, in other words, not off some kind of translation table that attaches specific meanings to specific sounds (heaven forbid). Just as interesting is 'das' which becomes something more like 'dash' as she moves her tongue back from the 's' position to 'sh' so that not only are the vowels the same, 'das Grab', but so are

²⁹ Originally issued as Decca 52832, Odeon O-7652, Parlophon P 9663, and Parlophon P 9871 (matrix 2-20717-2). At the time of writing a CD reissue was 'The Art of Meta Seinemeyer', Preiser PR 89402, disc 3, track 9.

³⁰ Battle & Levine, track 15.

the consonant positions for 'sh' and 'gr'. An unnatural but eerie frozen effect is produced by unnatural 's' sound and the unchanging mouth positions, as if the body were frozen into immobility with terror. In *Die Männer sind méchant*, we heard the voice evoking, copying the sounds of anger and imminent violence; here it reproduces the effects of fear on the voice.

Similarly, Lotte Lehmann in her 1941 recording uses upward pitch-sweeps very strongly for terror:³¹ '/Und /finster die /Nacht, wie /das /Grab' where '/' indicates a very fast but deep sweep up to the notated pitch. This comes straight from a highly dramatic reading of the text as speech. And indeed an LP recording of Lehmann reading part of 'Der Wegweiser' from *Winterreise* just a few years later shows this very clearly.³²

Und ich wandre sonder Maßen,	And I wander on relentlessly,
/Ohne /Ruh', und /suche Ruh'.	Restless, yet seeking rest.

And incidentally her 'Ohne Ruh' is very much like Battle's singing of that very similar text phrase in *Rastlöse Liebe*. When *Battle* does this she's drawing on an older tradition, then, or perhaps on a continuing operatic tradition, because Lieder singers stopped using sweeps in this way in the 1960s and 70s. Elly Ameling in 1975 doesn't at all.³³ What she does is to reduce her tone by using no vibrato, especially on 'wie'. Where does this immobile tone come from? What does it evoke or call on for its meaning? I'm not sure: possibly the cinema with its sleepwalkers and zombies: it's a different kind of reaction to horror, a frozen automaton response, the cultural and perceptual origins of which need further investigation.

Even less differentiated, one might think, is Arleen Auger, recorded with fortepiano accompaniment in 1990 and thus ostensibly working within the so-called 'historically informed performance' movement.³⁴ On the face of it her performance is exceptionally uninflected. But one has to read this in the context of its time and aesthetic, the period of

³¹ Originally issued as Columbia 71509-D and Columbia LOX 654 (matrix XCO 30013-1), reissued on 'Lotte Lehmann: Schubert', Dante Lys 231-234, disc 3, track 14 (issued 1997).

³² 'Lotte Lehmann Reading German Lyric Poems', Caedmon TC-1072 (issued 1957), side 2. There is a copy in the British Library Sound Archive, call number 1LP0062812.

³³ Elly Ameling accompanied by Dalton Baldwin, 'Schubert Lieder', Philips 464 334-2, disc 2, track 9 (recorded August 1975, this issue 1999).

³⁴ Arleen Auger, accompanied by Lambert Orkis, 'Schubert Lieder', Virgin Classics 5 61457 2, disc 1, track 21 (recorded March 1990, this issue 1998).

authenticity and deliberately non-interventionist readings of musical texts. In this context its deviations are very telling, tiny in absolute terms, but large relatively. Auger sings ‘und finster die NachTT, und FFinstER DIE Nacht’, where the very sudden loud TT is unexpected, an analogy to making you jump, the FF is started with an almost vertical sweep (a similar effect, but the different means increases the sense of unease), and the ‘-er die’ is wobbly with fright, as a spectrogram shows clearly as well.

These examples differ in the ways in which they represent something like fear; but they differ mainly because they belong within different period performance styles. At any time, singers find ways of representing sounds from life, not just using the notes provided by the composer, but also using the vocabulary of expressive gestures that constitute the performance style of their time or their generation. Fear is a clear example, but there are other kinds of references that can be made. As a completely opposite example, let’s consider infant-directed vocalisation. Elena Gerhardt’s 1928 recording of Schubert’s lullaby *Schlaflied* shows precisely that, especially in its exceptionally broad use of portamento.³⁵ Stephen Malloch has analysed the vocalisations used by mothers in talking to babies, including U-shaped pitch curves which Gerhardt may very well be modelling here. And the Papouseks have noted how falling contours and slowing tempos in infant-directed speech promote sleep or calm.³⁶ Schubert provided materials that would assist a singer in producing this effect, but Gerhardt adds to them with a use of portamento that greatly intensifies it. A singer today, of course, wouldn’t use portamento in this way; but Gerhardt simply draws on the most powerfully appropriate gesture that’s available in the performance style of her time and applies it.

Musical gestures can be borrowed from vocal communication, then, but from other experiences too. In fact, I suggest, anything that can be expressed through changes in shape

³⁵ Originally issued as HMV D 1460 (matrix Cc12964-2). Available for download from http://www.kcl.ac.uk/kis/schools/hums/music/ksa/ksa_sound.html and www.charm.kcl.ac.uk (follow the links to Schubert song sound clips). Also available as a CD transfer on Dutton, CDBP9727, track 4 (issued 2004). A more detailed study of this performance and its implications for our understanding of portamento is currently in preparation (late 2005).

³⁶ Stephen Malloch, ‘Mothers and Infants and Communicative Musicality’, *Musicae Scientiae*, special issue (1999/2000), 29-57. Hanus Papousek and Mechthild Papousek, ‘Musical elements in the infant’s vocalization: their significance for communication, cognition, and creativity’, in ed. Lewis P. Lipsitt and Carolyn K. Rovee-Collier, *Advances in Infancy Research* (Norwood, NJ: Ablex, 1981), I, 164-224. I am grateful to Anna Metcalfe for these references.

over time can be expressed through music. Which is why music can evoke the sea, but not a road, or pain, but not hunger. Pain has a shape – and so a sharply increasing dissonance can model it well – hunger doesn't. The sea can be modelled as waves, a road can't: it doesn't move. These are very crude examples, but it's obvious that affective science is going to have a great deal to discuss with musicians, because music can model the sensation of emotions.

The last recording from which I should like to draw is of Janet Baker's live performance of Schubert's *Abendstern* (Evening Star) sung at a BBC concert in 1980.³⁷ The star is the star of love, who, says the text, sows love in others but receives none herself, she remains alone, sad and silent. Baker's performance is marked by restrained tone, the narrowing of her usual vibrato, unusually even amplitude level, and very plain delivery, all are there to evoke conventional associations with icy stars (traditionally we see them as cold, like snowflakes, rather than as suns), in this case frozen by lack of love. So a whole complex of conventional images is evoked by the sounds she makes. Especially still is her singing of

Ich bin der Liebe treuer Stern;	I am the faithful star of love;
Sie halten sich von Liebe fern.	the other stars stay far away.

And within this, the emotive word 'Liebe' is treated with restraint, of course, a little portamento, a widening of vibrato, but within tightly constrained limits. What interests me, though, is its shape. The word gets louder, and that brings with it a spectrum change: the sound brightens as upper harmonics come in, the 'ie' getting more metallic, and the pitch rises very slightly, then returns. So as well as the very faint swoop that marks the note at its onset as special, the progress of the note has a shape to it. It's a complex event that lends the word and the bare musical material much greater weight than the notes around it. Yet at the same time, because the changes in the most easily perceived dimensions – vibrato, amplitude – are very small, and the biggest change is in a dimension to which we don't normally attend closely – the timbre – and because even that happens gradually, this rather special note does not significantly disrupt the coherence of the line. This example emphasises how the quality of a musical sound can have a role to play in structuring a performance. But what exactly is it doing here to be expressive of love?

Many of the things I've been talking about, and especially the musical analogues – or signs – of emotional states, need to be understood in the light of the way the brain processes sounds and the mind reacts to them. These musical and extra-musical sounds are meaningfully

related not just as a system of symbolic relations or analogies but as the equivalent output of related systems in the mind. In other words, something has been cross-domain mapped, or blended, between feeling and sound. And clearly the notion of cross-domain mapping is crucial to this whole subject. Related to this, inevitably, are the evolutionary reasons for the close link between sound and emotion. In the interests of survival sounds require the fastest possible response – the opportunity or threat assessment has to be made as fast as the physical system, via natural selection, can be engineered to make it. The immediate reaction to sound is therefore emotional, not intellectual. One feels something about the sound and what it means first of all, and has time to think about it later. And what's very striking about the way we respond to music, even now, is that emotional reaction is fundamental for all listeners except those academic musicians who have been trained to side-step it and to treat sound analytically from the start. For everyone else music triggers or represents or mimics or evokes or models emotions (which of these it does may well vary according to other conditions, as Scherer and Zentner have suggested,³⁸ but is in any case a question for further, scientific research). Of course, music does these things for academics too when we let it. One of the implications of this view, and of the way I'm looking at performance, is that we should let it much more than we do; that we would gain another dimension, not revert to childishness in our response to music. Academics fear the emotions, because emotions are thought to be subjective and irrational, unlike scholars. But emotions are what music speaks to before anything else; they are inextricably bound up with the ways it works.

So, coming back to Janet Baker's *Abendstern*, we can see that the structure that this piece has in this performance, or in any performance, is not just a formal structure but also an emotional structure. And in fact Baker's particular treatment of 'Liebe' is as good an example as any of how emotional structure works. Why is a shape that increases and then decreases – which this does all its dimensions – evocative of deep emotion? It's an absolutely fundamental shape in music, both in performance and in composition. Most gestures, whether in composition or in performance, get larger and then smaller; they get louder and then softer; they go up and then down, as in most melodic phrases; they go away and then come back, the basis of tonality as a process; or more rarely they do the opposite, getting softer and then louder, lower then higher. It's evident that they are invoking a basic shape in our experience of emotion. Music can do this because it happens in time, just as

³⁷ Janet Baker accompanied by Geoffrey Parsons, 'Baker: Schubert Lieder', BBC Legends BBCL 4070-2, track 14 (recorded September 1980, issued 2001).

³⁸ Klaus R. Scherer and Marcel R. Zentner, 'Emotional effects of music: production rules', in Juslin & Sloboda, 361-92 at 361.

emotions happen in time, and because, like emotions, it is a process that starts from nothing, moves through something, and returns to nothing at its end. Emotions, of which love is simply a conveniently strong example – but any other would do: anger, surprise, fear, curiosity, excitement, misery – every one is a temporary phenomenon with a shape; it grows, peaks and then subsides. Music makes an art out of that by timing it right, so that shapes in different dimensions (harmony, melody, rhythm, texture, but also colour, speed, amplitude and numerous other aspects of performance) coincide or interact in suggestive and interesting ways. It follows that research into these emotional shapes, into what emotions feel like, could be highly relevant to studies of musical expressivity. There've been a number of attempts in the past, from Langer through to Clynes,³⁹ but the rapid development of affective science offers new opportunities.

Let me summarise in a few sentences the suggestions I've made that point towards a wider hypothesis. Scores are notations of musical materials that performers and listeners together make into music. Performances should be the focus for empirical work on musical expressivity. Performers bring meaning to musical materials by referring to the sounds or shapes or other qualities of experiences we know from life. Listeners recognise these references and experience the music as a representation in the musical domain of feelings that they recognise. (I would add that the process of recognition may well work in the same way as Neumann and Strack have shown for mood contagion: sounds associated with moods cause the automatic transfer of mood between persons.⁴⁰) By representing sounds we associate with feelings, music similarly causes the automatic evocation of mood for listeners.

One final issue has to be addressed, if only very briefly. All my examples have been of music with text, and the text obviously suggests the emotion the performer is supposed to express and the listener to perceive. I make four points about this. 1) By starting with texted examples we're in much better position to say how musicians understand the relationship between sounds and emotion. 2) Expressivity in instrumental playing uses exactly the same performance gestures as expressivity in singing, and it's inconceivable that they work in different ways. 3) Expressivity in instrumental performance can therefore be understood as an adaptation of vocal expression. 4) In general, and even in singing except where the text is absolutely specific, performance gestures should not be pinned down to a single specific

³⁹ For an outstandingly good survey see Gabrielsson & Juslin, 'Emotional expression', cited above.

⁴⁰ Roland Neumann and Fritz Strack, "'Mood Contagion": the automatic transfer of mood between persons', *Journal of Personality and Social Psychology* 79 (2000) 211-223.

meaning. Musical expressivity isn't fixed to that extent. Music is expressive but only occasionally of something specific. We should be more concerned to understand how music is expressive than to try to show what it must be expressive of.

To conclude, it seems to me that by turning away, to some extent, from history and analysis of scores, and by starting to ask questions about the history of recorded performance, musicology has found itself facing exactly the same questions as psychologists of music perception are facing, and is finding answers in exactly the same places. Our research aims are essentially identical. Our methodologies, however, are not. Perhaps they should be. Perhaps musicologists should become scientists and leave the speculation and guesswork of the humanities behind. But I'm not sure that that would help science. What I think musicology can offer is, first of all, easier access to the best-informed understandings of what it's like to make music with the body and in the mind, and secondly a freedom to guess that will generate more fruitful hypotheses for scientists to test.