

The Multimodal Experience of Art

Bence Nanay

The aim of this paper is to argue that our experience of artworks is normally multimodal. It is the result of perceptual processing in more than one sense modality. In other words, multimodal experience of art is not the exception; it is the rule. I use the example of music in order to demonstrate the various ways in which the visual sense modality influences the auditory processing of music and conclude that this should make us look more closely at our practices of engaging with artworks.

The Multimodality of Perception

One of the most important new directions in contemporary philosophy of perception is to consider the far-reaching consequences of the recent body of literature on the multimodality of perception. There is a lot of recent empirical evidence that multimodal perception is the norm and not the exception—our sense modalities interact in a variety of ways.¹ It is difficult to overstate the importance of these findings for some of the classic debates in the philosophy of perception. One quick example. The question about the individuation of the senses, for example, as it has been raised by philosophers of perception, presupposes that the sense modalities are unimodal. If, as we now know, sense modalities are not unimodal, we need to re-evaluate this debate.

If our perceptual experiences are typically multimodal, then we also need to re-evaluate some questions in aesthetics. It has been assumed both in philosophy of perception and in aesthetics that, say, visual experience is unimodal: it is not influenced by what goes on in the other sense modalities. But this turns out to be false: our visual experience very much depends on what goes on in our other sense modalities.² And if this is true, then those questions in aesthetics that are about our perceptual experience of artworks need to take these new insights into consideration. This article is intended to be a first step into that direction.

It is important to note that I am not suggesting that we should directly import the results of empirical psychology to aesthetics. The direct application of empirical results in aesthetics can, and very often does, go terribly wrong. What I suggest is that aesthetics should take some new paradigms of philosophy of perception seriously. The specific paradigm I am interested in here, the paradigm of multimodality, is based on a large body of empirical research. However, my aim is not to urge an empirical turn in aesthetics, but

1 See Casey O'Callaghan, 'Object Perception: Vision and Audition', *Philosophy Compass* 3 (2008), 803–29, and 'Seeing What You Hear: Cross-modal Illusions and Perception', *Philosophical Issues* 18 (2008), 316–38, for philosophical overviews.

2 See C. Spence and J. Driver (eds), *Crossmodal Space and Crossmodal Attention* (Oxford: OUP, 2004); P. Bertelson and B. de Gelder, 'The Psychology of Multimodal Perception', in C. Spence and J. Driver (eds), *Crossmodal Space and Crossmodal Attention* (Oxford: OUP, 2004), 141–77, for summaries.

to urge a turn in aesthetics towards philosophy of perception, and this sometimes entails a turn towards empirically informed philosophy of perception.³

Let us see what the multimodality of perception amounts to. Information in one sense modality can influence the information processing in another sense modality. This influence can happen at a very early stage of perceptual processing (often in the primary visual cortex in the case of vision).⁴ A simple and neat example of the multimodality of perception is ventriloquism.

Ventriloquism is one of the prime examples of what is known as ‘crossmodal illusions’, where different sense modalities give us conflicting information about the world and this conflict is resolved in our overall experience. In the case of ventriloquism, the visual sense modality tells our perceptual system that the sound comes from the dummy, whereas the auditory sense modality tells our perceptual system that it comes from the ventriloquist. The way this conflict is resolved is that we experience the voices as coming from the dummy and not from the ventriloquist (where they actually come from).⁵ In this case, as in most (not all) cases of crossmodal illusions, vision wins out: it influences our audition and not the other way round.

Another demonstration for crossmodal effects where vision trumps audition is the McGurk effect: the visual stimulus of the speaker’s mouth alters the auditory experience of the sound we hear the speaker make: the auditory stimulus is the same (say, a ‘b’ sound), but depending on the visual stimulus of the speaker’s mouth (whether she makes lip movements that correspond to the utterance of the ‘b’ or of the ‘v’ sound), our auditory experience will be different. The visual information about the speaker’s mouth appearing to utter a ‘v’ sound can overwrite the auditory information about the ‘b’ sound.⁶

But there are more surprising examples of multimodal perception: if there is a flash in your visual scene and you hear two beeps while the flash lasts, you experience it as two flashes.⁷ This is one of the not so many examples where vision does not trump audition: the two beeps in our auditory sense modality influence the processing of the one flash in our visual sense modality; and, as a result, our visual experience is as of two flashes.

3 Bence Nanay, *Aesthetics as Philosophy of Perception* (Oxford: OUP, forthcoming), and ‘Aesthetics as a Subdiscipline of Philosophy of Perception’, in Gregory Currie, Matthew Kieran, and Aaron Meskin (eds), *Aesthetics and the Sciences of the Mind* (Oxford: OUP, forthcoming).

4 See, for example, S. Watkins *et al.*, ‘Sound Alters Activity in Human V1 in Association with Illusory Visual Perception’, *NeuroImage* 31 (2006), 1247–56; R. Sekuler, A. B. Sekuler, and R. Lau, ‘Sound Alters Visual Motion Perception’, *Nature* 285 (1997), doi:10.1038/385308a0; J. Vroomen, P. Bertelson, and B. de Gelder, ‘Auditory-Visual Spatial Interactions: Automatic Versus Intentional Components’, in B. de Gelder, E. de Haan, and C. Heywood (eds), *Out of Mind* (Oxford: OUP, 2001), 140–50.

5 P. Bertelson, ‘Ventriloquism: A Case of Cross-modal Perceptual Grouping’, in G. Aschersleben, T. Bachmann, and J. Müssele (eds), *Cognitive Contributions to the Perception of Spatial and Temporal Events* (Amsterdam: Elsevier 1999), 347–62. See also David Goldblatt, *Art and Ventriloquism* (London: Routledge, 2006).

6 H. McGurk and J. MacDonald, ‘Hearing Lips and Seeing Voices’, *Nature* 264 (1976), 746–8.

7 L. Shams, Y. Kamitani, and S. Shimojo, ‘What You See Is What You Hear’, *Nature* 408 (2000), doi:10.1038/35048669.

The Multimodality of the Experience of Artworks

If perception in general is multimodal, then it would be surprising if the perception of artworks were not multimodal. And there is some (not an overwhelming amount of) work on the multimodality of our experience of at least some aspects of some kinds of artworks. But these are mainly limited to our experience of music and, more precisely, to our experience of the expressiveness of music. More specifically, they demonstrate that visual stimuli play an important role in our aesthetic appreciation of the expressiveness of musical performances.⁸

The aim of this paper is to explore some more general consequences of the multimodality of perception for the way we should think about the experience of art. The experience of art is genuinely multimodal: the visual stimulus can and does modify and influence what we take to be our auditory experience (say, of music) and the auditory stimulus can and does modify and influence what we take to be our visual experience (say, of pictures).

In order to focus the discussion of the multimodality of art, I need to put aside some rather trivial and uninteresting cases of multimodal experience. I am in a museum but the loud tour guide in the next room keeps distracting me. Is this a genuine multimodal experience? No, not really. The auditory sense modality influences my overall experience: it makes me annoyed. But it does not influence my visual experience, at least not directly: as a result of being annoyed, I may find it more difficult to focus on the painting in front of me, but the auditory sense modality does not change the way my visual sense modality processes the visual features of the painting.

What I mean by the multimodal experience of art is not just that we use more than one sense modality when engaging with artworks—we, admittedly, often do so, for example, in the opera or when watching a film. One's experience of art is multimodal if information in one sense modality influences not merely one's overall experience but one's perceptual experience in another sense modality.

I will mainly focus on the multimodal experience of music in this article. The reason for this is twofold. First, there has been more research into how our experience of music is influenced by the other sense modalities, especially by vision than about the multimodality of any of the other arts.⁹

-
- 8 V. Bergeron and D. Lopes, 'Hearing and seeing musical expression', *Philosophy and Phenomenological Research* 78 (2009), 1–16; Bradley W. Vines *et al.*, "Cross-modal Interactions in the Perception of Musical Performance," *Cognition* 101 (2006), 80–113, and 'Dimensions of Emotion in Expressive Musical Performance', *Annals of the New York Academy of Sciences* 1060 (2005), 462–66; and Jane W. Davidson, 'Visual Perception of Performance Manner in the Movements of Solo Musicians', *Psychology of Music* 21 (1993), 103–12.
- 9 Catherine Chapados and Daniel J. Levitin, 'Cross-modal Interactions in the Experience of Musical Performances: Physiological Correlates', *Cognition* 108 (2008), 639–51; Bradley W. Vines *et al.*, 'Music to My Eyes: Cross-modal Interactions in the Perception of Emotions in Musical Performance', *Cognition* 118 (2011), 157–70; W. F. Thompson and F. A. Russo, 'Facing the Music', *Psychological Science* 18 (2007), 756–7; J. Davidson, 'Visual Perception of Performance Manner in the Movements of Solo Musicians', *Psychology of Music* 21 (1993), 103–13; V. Best *et al.*, 'The Influence of Spatial Separation on Divided Listening', *Journal of the Acoustical Society of America* 120 (2006), 1506–16; W. F. Thompson, P. Graham, and F. A. Russo, 'Seeing Music Performance: Visual Influences on Perception and Experience', *Semiotica* 156 (2005), 177–201; W. F. Thompson, F. Russo, and L. Quinto, 'Audio-visual Integration of Emotional Cues in Song', *Cognition and Emotion* 22 (2008), 1457–70;

Second, and more importantly, in the case of the experience of music, we get systematic and aesthetically interesting influences from the visual sense modality, whereas in the other way round, much of the multimodality of our experience of the visual arts could be dismissed as mere disturbances on our otherwise pure and purely visual experience (with the important exception of film music, see below).

I explore six important forms of this interaction in the next section—six case studies of the importance of the multimodal nature of our musical experience: (a) highlighting and emphasizing musical form, (b) serving as a counterpoint for musical form, (c) obscuring musical form, (d) highlighting and emphasizing the expressive content of music, (e) serving as a counterpoint for the expressive content of music, and (f) obscuring the expressive content of music.

A Case Study: Multimodal Influences on Music

To keep the discussion of the various kinds of multimodal influences on our experience of music as simple and straightforward as possible, I will make a distinction between musical form and expressive content. I do not mean this to be an absolute distinction, nor would I want to commit to any kind of duality of content and form in general. The reason why I need to keep apart these two general categories is to make the typology of the multimodal influences on the experience of music easier to handle. Further, I do not want to exclude the possibility that by influencing our experience of musical form, some multimodal effects also influence our experience of the expressive content or vice versa. But for the sake of simplicity I will discuss these two general forms of multimodal influences on our experience of music separately.

Highlighting and Emphasizing Musical Form

The most obvious example of information in the visual sense modality highlighting or emphasizing the auditory experience of musical form is the conductor's hand movements that emphasize and highlight certain formal elements of music. Nikolaus Harnoncourt's conducting, with his usually economical movements that only burst into gestures at formally significant points, provides an excellent example. Most of the time, he merely dictates the rhythm—like many other conductors. But occasionally, when something important is happening in the score, he suddenly bursts into an energetic gesture that draws our (visual) attention to what is going on in the musical score at that moment, thereby making the musical form more salient.

M. Schutz and S. Lipscomb, 'Hearing Gestures, Seeing Music: Vision Influences Perceived Tone Duration', *Perception* 36 (2007), 888–97; M. Schutz, 'Seeing Music? What Musicians Need to Know About Vision', *Empirical Musicology Review* 3 (2008), 83–108; S. Dahl and A. Friberg, 'Visual Perception of Expressiveness in Musicians' Body Movements', *Music Perception* 24 (2007), 433–54; and M. Broughton, C. Stevens, and S. Malloch, 'Music, Movement and Marimba: An Investigation of the Role of Movement and Gesture in Communicating Musical Expression to an Audience', *Psychology of Music* 37 (2009), 137–53.

Other examples where vision highlights and emphasizes musical form includes some ballet and modern dance choreographies, such as by Mark Morris or Jiri Kylian. Both of these choreographers tend to adjust their choreography to the music in a (sometimes almost comically) synchronous manner. Take Jiri Kylian's choreography 'Birthday' for the Nederlands Dans Theater (2006) that uses the music of Mozart's overture of *Le nozze di Figaro*. Everything the two dancers do in the kitchen (sneeze, cut the dough, break eggs, etc.) is synchronous with the most important musical features—this often leads to comical effects. This choreography makes the musical features that are accompanied by synchronous visual impulses much more salient.¹⁰

Serving as a Counterpoint for Musical Form

But vision does not always serve to emphasize and highlight musical form. Often, it does the exact opposite: it serves as a counterpoint. Take the famous performance of Rameau's *Les Indes galantes* by Les Arts Florissants, conducted by William Christie and choreographed by Blanca Li and Andrei Serban (2004, Opéra national de Paris). The choreography of the duet 'Forêts paisibles' in the last act between Zima and Adario involves very pointed visual gestures against the beat, which makes our multimodal experience of this performance of the duet shift time signature. We hear it as having the time signature of 4/4 instead of the original *alla breve* time signature (2/2) as prescribed in Rameau's score. Here what we see (gestures against the beat) makes us experience the formal properties of the music differently.

To turn to modern dance, some of Pina Bausch's choreographies use the same effect. At the beginning of her *Café Müller* (1978, Tanztheater Wuppertal), the woman's movements almost always seem to be the exact opposite of what is happening in the musical score (of 'O let me weep' from Purcell's *The Fairy Queen*). She stands still for a long time and then suddenly, when there is a lull in the music, starts running; she makes frantic complicated gestures while the music is slower, and hardly moves when the music gets faster. The same applies to Bausch's choreography for Gershwin's 'The Man I Love' in her *Nelken* (1982, Tanztheater Wuppertal), where the man's gestures are supposed to express the same meaning as the song's lyrics, but their timing is almost always against the beat. In this interesting example, the auditory experience of both the musical form and the expressive content is influenced by visual effects.

Obscuring Musical Form

I considered examples where information in the visual sense modality influences our auditory experience of musical form either by highlighting it or by counterpointing it. A third kind of multimodal influence on our auditory experience of musical form is more complex and more sophisticated than these two earlier kinds and it can be aesthetically very

¹⁰ See also C. L. Krumhansl and D. L. Schenck, 'Can Dance Reflect the Structural and Expressive Qualities of Music? A Perceptual Experiment on Balanchine's Choreography of Mozart's Divertimento No. 15', *Musicae Scientiae* 1 (1997), 63–85.

significant. Sometimes the visual sense modality does not add to our auditory experience of musical form (by reinforcement or by counterpoint), but rather takes away from it. In other words, sometimes the visual sense modality makes our auditory experience of the musical form more ambiguous—it obscures musical form.

The best examples for this kind of multimodal effect come from modern dance. The clearest cases are choreographies of Trisha Brown, who very explicitly attempts to make her choreography as asynchronous with the music as possible. Take her choreography for Rameau's *Pygmalion* (with Les Arts Florissants, Festival d'Aix-en-Provence, 2010). The dancers' movements very deliberately avoid either emphasizing or providing counterpoint for the music—they form a parallel, but independent perceptual stimulus. The effect is some kind of dislodgement of the musical forms, including the rhythm and the metric—they become much less clearly defined as a result of the visual experience of the choreography.

Highlighting and Emphasizing the Expressive Content of Music

So much for musical form. My examples of the multimodal influences of our experience of the expressive content of music come from film music. Film music is a good place to start when looking for multimodal effects in our experience of art, but while there are many philosophically sensitive analyses of film music and the experience thereof, the emphasis is invariably on how the addition of music changes our visual experience of what happens on the screen.¹¹ This is undoubtedly an important example of the multimodal experience of film and it should play an important role in any analysis of how the auditory sense modality influences the visual experience of art, but the case studies I am focusing on in this article are about crossmodal influences of the opposite direction: visual influences on our auditory experience of music.¹² And, alas, this aspect of film music has largely been ignored.

The expressive content of film music can undoubtedly influence the visual experience of what is on the screen. The question I would like to examine is whether the visual experience of what is on the screen can influence the expressive content of film music (while acknowledging that the influence can be, and probably very often is, bidirectional).

A good test case for this is if we consider films that use the same music for accompanying different images. At the beginning of Tarkovsky's *Sacrifice* (1986), we hear the 'Erbarne dich' aria from Bach's *St Matthew's Passion* while we see details of Leonardo's

11 Jerrold Levinson, 'Film Music and Narrative Agency', in D. Bordwell and N. Carroll (eds), *Post-theory: Reconstructing Film Studies* (Madison, WI: University of Wisconsin Press, 1996), 254–88, and 'Sound in Film: Design versus Commentary', unpublished manuscript; David Bordwell and Kristin Thompson, 'Fundamental Aesthetics of Sound in Cinema', in *Film Art: An Introduction* (4th edn, New York: Knopf, 1993); Claudia Gorbman, *Unheard Melodies: Narrative Film Music* (Bloomington, IN: Indiana University Press, 1987); George Wilson, *Narration in Light* (Baltimore, MD: Johns Hopkins University Press, 1986); Kathryn Kalinak, *Settling the Score: Music and the Classical Hollywood Film* (Madison, WI: University of Wisconsin Press, 1992); and Elisabeth Weis and John Belton (eds), *Film Sound* (New York: Columbia University Press, 1985).

12 These crossmodal influences are of course not necessarily unidirectional: if our auditory experience changes the way we perceive visually, then this visual experience can in turn influence our auditory perception.

Adoration of the Magi. This is the first shot of the film, so we have no previous knowledge of either the characters or the narrative. Contrast this auditory experience of the aria with the one in Pasolini's *Il Vangelo secondo Matteo* (1964), where we hear the same aria during Peter's denial of Christ and while he realizes what he had just done. And, to use a third example, in Carlos Reygadas' *Japon* (2002), the same aria is heard after the main character's failed suicide in the rain in a long aerial shot of him and a dead horse. Our experience of the same aria in these three films couldn't be more different.

Here is another example: Godard uses Mozart's Concerto for Clarinet and Orchestra in A Major (K622) both in *Masculin féminin* (1966) and in *Breathless* (1959). In *Masculin féminin*, it follows, somewhat surprisingly (especially given that at that time Godard very rarely used non-diegetical music) a cheerful conversation in the laundromat between two young men about Bob Dylan, Vietnams, and the revolution. In *Breathless*, he uses it diegetically at the point of the narrative where Patricia decides to report Michel, her lover, to the police.

The phenomenal character of our auditory experience of the same pieces of music (admittedly, in different recordings in the case of 'Erbarne dich') is very different, as a result of the visual stimulus that accompanies them. Importantly, it is our experience of the expressive content of the music that differs, not (or not primarily) our experience of the musical form. This is a good example of the second general form the multimodal experience of music takes: the visual influences the experience of the expressive content of music.

But this crossmodal influence can take various different forms. The first, and in some ways the simplest, case is where the images emphasize the expressive content of the accompanying music. The use of 'Erbarne dich' in Pasolini's film is a good example: the images of Peter looking heartbroken and starting to cry after his denial of Christ emphasize and highlight the aria's emotional and tragic overtones (in a way that the images of the details of Leonardo's painting do not).

Another good example of this way of using film music is from another Pasolini film, *Mamma Roma* (1962). In the last scene, where Ettore dies and where his mother attempts to commit suicide, we hear the largo movement of Vivaldi's Concerto in D minor for Viola d'amore and Lute (RV 540)—one of the most tragic of Vivaldi's slow movements. The images make our experience of this music even more tragic. Pasolini's other (early) films are also full of this way of using film music (especially *Accattone*, 1961 and *Il Vangelo secondo Matteo*, 1964).

The expressive content that influences our experience of music is not necessarily negative (as the examples above may suggest). Truffaut's *Antoine et Colette* (1962) starts with a scene where the young Antoine Doinel listens to the air movement of Bach's Third Orchestral Suite and steps out to his balcony looking at the view of Paris triumphantly. And this makes the music even more triumphant. In István Szabó's *Apa* (1966), there is a scene, accompanied by the restless and energetic first movement of Bach's Harpsichord Concerto No. 1 in D minor (BWV 1052), where after the war, an old streetcar is pushed through the streets of Budapest, giving hope and a sense of a new beginning to the people. These images, again, make the music even more restless and energetic.

The famous levitation scene of Tarkovsky's *Solaris* (1972) uses the same effect, in a more complex form: we hear a rendering of Bach's 'Ich ruf zu dir' (BWV 639) while we see,

besides the levitation of the couple, details of Breughel's *Hunters in the Snow* and a large body of water swirling in slow motion. It is difficult to pin down what the expressive content of these images are and how exactly they influence our experience of the Bach piece, but there is some general parallel between the serene and solemn images that make the music even more serene and solemn. This is also the general pattern for the way in which Tarkovsky uses music in his other films, for example, in *Mirror* (1975).

Serving as a Counterpoint for the Expressive Content of Music

But the expressive content of the images is not always congruent with that of the music. An interesting way of using film music is to pit the expressive content of the images against the expressive content of the music. We have already seen one example of this from Godard's *Breathless*, where we hear the first movement of Mozart's Concerto for Clarinet and Orchestra in A Major (K622) while seeing Patricia's face in close up, while she is contemplating reporting Michel to the police. The music is not at all tragic or melodramatic or ominous—it is (in the recording Godard uses) almost cheerful, in sharp contrast with what is happening on the screen. Our auditory experience of the music is overshadowed by the visual experience of seeing Patricia's betrayal.

Another example from Godard: In *Masculin féminin*, he uses the second movement of the same piece, also in a diegetic context. This slow movement is emotional and moving in an almost romantic manner (especially in the recording Godard uses), but the scene itself has nothing of the emotional effects the music suggests. Paul puts on the record, pontificating, in a somewhat ridiculous manner, about the orchestra in the background. This makes our experience of the music much less moving than it would be without the images.

The ending of Robert Bresson's *Mouchette* (1967) has the opposite effect. The music is a very upbeat excerpt from Monteverdi's *Magnificat*, while the scene is about the desperate suicide of the thirteen-year-old Mouchette. These images work against the music, making our experience of the Monteverdi piece much less energetic and upbeat.

Another, somewhat different, way of taking advantage of the contrast between the expressive content of images and sound comes from Godard's *Le Mépris* (1963). During the opening, very long shot, we hear Georges Delerue's extremely sentimental and ominous score, which would serve as the ideal film music in a traditional crime fiction film. But the shot is of a sunny Italian street with a small crew shooting a film scene—nothing ominous, nothing sentimental. Again, our experience of the otherwise emotionally loaded music is made less emotional because of the images.

To use a less highbrow example, the comic effects of the conflict between image and sound are widespread in the popular media. In an episode of the American sitcom *How I Met Your Mother* (season 4, episode 2), one of the characters describes the best burger he has ever had with the musical accompaniment of the adagio movement of Mozart's 'Gran Partita' (Serenade No. 10 in B Flat Major, K 361). The same effect is also used in a number of Monty Python sketches.

In the examples above, the expressive contents of the images and of the music are in conflict. But sometimes we also get a counterpoint effect when there is no such conflict, but the images nonetheless do weaken the expressive content of the music. Here is an

example from the beginning of Louis Malle's *Viva Maria!* The music here is also by Georges Delerue (the iconic composer for a number of nouvelle vague films), an intentionally cheesy piece about the adventurous childhood of the young Maria, accompanied by even cheesier images. At one point, when, after the death of her father Maria is running away from the English troops, she is running across a suspension bridge with the sunset and the dramatic Central American landscape in the background, while the music bursts out in a particularly sentimental tirade, the cheesiness is just too overwhelming—the cheesiness of the images makes the cheesy music downright ridiculous.

Obscuring the Expressive Content of Music

Finally, sometimes the expressive content of the images are neither congruent with, nor diametrically opposed to, the expressive content of the accompanying music. The most important examples for this are ones where the images do not have clear expressive content. Therefore, what we see does not have a very obvious and straightforward influence on our experience of the music. But this does not mean that it has no influence. It is just difficult to have a simple model for what this influence is: it is not reinforcement and not counterpoint either.

Here is an example: the use of Bach in Godard's *Je vous salue Marie* (1985)—such as in the scene where we see the main character standing in a window playing with a Rubik cube. These images do not emphasize the expressive content of the Bach piece, nor do they work actively against it. Rather, as we have difficulty interpreting the expressive content of the images, it has an effect on our auditory experience of the music that makes this experience have a less definite expressive content. In short, it makes this expressive content more ambiguous. Godard also uses the same effect in his *Armide* (1987), where he shows images of bodybuilders to obscure the expressive content of the accompanying Lully opera.

Another clear example of this way of using music is from some of Bunuel's early films. In *Un Chien Andalou* (1929), while it is a silent film, Bunuel insisted on using the 'Liebestod' from Wagner's *Tristan und Isolde*. The images are famously difficult to make sense of and this obscures the expressive content of the otherwise somewhat romantic Wagner piece. The same goes for the music in his *L'Age d'or* (1930)—for example the famous opening scene with the scorpions accompanied by Mendelssohn's *Hebrides* overture (Op. 26) or the party scene accompanied by the opening movement of Schubert's *Unfinished Symphony*.

A final example. At the end of both *Through the Olive Trees* (1994) and *Life and Nothing More* (1992), Abbas Kiarostami uses baroque music (quite atypical for the director) in the very last scene. In both films, the last scene consists in one long shot, with the characters very far away and it is impossible to tell what they are doing exactly. In the case of *Through the Olive Trees*, we see a man, Hossein, running on a field very far away to catch Tahereh, whom he wants to marry. We can see only two white dots far away as one of them catches up with the other and then runs back. The accompanying music is the allegro giusto movement of Cimarosa's *Oboe Concerto in C Major*—a bittersweet melody the expressive content of which becomes obscured by the fact that we have no idea what happens between Hossein and Tahereh so far away from the camera. The ending of *Life*

and Nothing More has the very same structure: we see a conversation very far away from the camera, not knowing what it is about. The auditory experience of the accompanying Vivaldi excerpts, as a result, is becoming somewhat detached.

There are many more examples of the way the auditory and the visual sense modalities interact to give us a genuinely multimodal experience of music. The ones I mentioned here I take to be paradigmatic, but they may not be the only ones. And some may question the way I analysed the specific examples (it could be argued that Godard, for example, always uses images and music in such a way that the images make our experience of the music more ambiguous—I would be very open to this suggestion).

But the general moral of all these examples is that visual and auditory information interacts in the overall multimodal experience on a variety of levels and in a variety of ways: both working against each other and in support of each other and both in terms of the form and the expressive content of music. Understanding how we (normally) experience music involves the understanding of these interactions.

The concept of perceptual experience has been playing a more and more important and diverse role in various debates in aesthetics. If it turns out, as I have argued, that the experience of artworks is not typically unimodal, this claim has a number of important consequences for these debates in aesthetics.

A quick example: an important debate in contemporary philosophy of music is about the authenticity of musical performances: about how musical works should be performed.¹³ Do we have to perform the musical work as it was intended to be performed by the composer? Do we have to perform it in such a way that the experience is comparable to that of the original performance? Whichever approach we take, if the experience of music is multimodal, then the debate about authenticity will look very different. If, for example, one argues that in the case of an authentic performance, the experience is comparable to that of the original, if this experience is multimodal, then this leads to very different criteria for evaluating the authenticity of performances.¹⁴

I explore a more general consequence of taking the experience of artworks to be multimodal in the last section.

Conclusion: Multimodality and the Question of What One Should Ignore

When we engage with an artwork, we invariably ignore some of its features and focus our attention on others. When admiring Petrus Christus's *Portrait of a Young Girl* (Gemeldegalerie, Berlin), we ignore the cracks in the paint and focus our attention on

13 P. Kivy, *Authenticities: Philosophical Reflections on Musical Performance* (Ithaca, NY: Cornell University Press, 1995); Jerrold Levinson, *Music, Art, and Metaphysics* (Ithaca, NY: Cornell University Press, 1990), 313–408; Stephen Davies, *Musical Works and Performances* (Oxford: OUP, 2001); James O. Young, 'The Concept of Authentic Performance', *BJA*, 28 (1988), 228–38; and J. Dodd, *Works of Music: An Essay in Ontology* (Oxford: OUP, 2007), 201–39.

14 See Bence Nanay, 'Musical Twofoldness', *Monist* 95 (2012), 607–24.

other features of the painting's surface. We abstract away from the cracks. When looking at a Romanesque church that was rebuilt in the Baroque era, we may try to ignore the Baroque elements in order to admire the medieval structure. Again, we are attempting to abstract away from some features of the artwork.

One big question is this: how do we know what properties of an artwork we should be paying attention to and what properties we should ignore or actively abstract away from. One answer is that we should ignore all those properties that the artist did not intend us to attend to. Petrus Christus clearly did not want anyone looking at his portrait to focus on the cracks in the paint—they were not there when he painted this small picture. Another way of distinguishing those properties that need to be attended to from those that need to be ignored would be to ask what would give us the highest degree of aesthetic experience/pleasure—perhaps independently of what the artist intended.

I do not intend to take sides in this debate here. But if we take the multimodality of the experience of art seriously, we can, and should, reframe this debate. My aim was to argue that our experience of works of art is multimodal. But if this is true, then we should not automatically ignore properties that are experienced in a sense modality different from the primary sense modality the artwork is experienced in. A tempting short cut for what properties to ignore is to shut out all the sense modalities that are not the ones the artwork is 'supposed to' be enjoyed with. Those members of the concert audience who close their eyes during the concert and often even during an opera production would provide a paradigmatic example of this strategy.

If it is true that our experience of artworks is genuinely multimodal, then this is a mistake. The aesthetically relevant properties of an artwork may be spread across sense modalities. Thus, if we are engaging with a piece of visual art, we should not automatically ignore anything auditory. Conversely, when we are listening to music, we should not necessarily abstract away from any visual stimulus. As we have seen above, some of this visual stimulus can colour and make more rewarding our experience. This is not to say that the aesthetically relevant properties of a musical work *always* include visual ones—but sometimes they do. And if we automatically discard anything visual when enjoying music, we will miss out on something aesthetically relevant and important.¹⁵

Bence Nanay
 Centre for Philosophical Psychology,
 University of Antwerp and Peterhouse,
 University of Cambridge
 bn206@cam.ac.uk

15 This work was supported by the EU FP7 CIG grant PCIG09-GA-2011-293818 and the FWO Odysseus grant G.0020.12N. I am grateful for comments by members of the audience at the Royal Musical Association Music and Philosophy Conference in July 2012 as well as an anonymous referee for this Journal.