

Timbre, Harmony, Orchestration, and Analysis, and Rautavaara’s Symphony no. 8 “The Journey”

Joshua W. Mills

Premises

My conclusion rests upon two premises:

1. The line that separates the perception of pitch from the perception of timbre is contextual, subjective, and fluid.
2. The notational conventions of Western musical practice depict timbre—and more importantly, timbral manipulation—only indirectly.

Since the second of these two premises is the more concrete, I shall begin there.

1 Timbre, Objectively

1.1 Notation and Timbre

Ever since Aristoxenus, the primary concern of music-theoretical discourse in our tradition has been pitch—modes, scales, canonic, counterpoint, harmony, set classes, geometric pitch space. This is mirrored in our notational system. A staff assumes, and then defines, a “pitch space” in which notes can exist. The staff is a two-dimensional space: the vertical axis represents pitch, and the horizontal axis, time. Musical motion, such as a melody, is represented as a temporal succession of discrete notes occurring within this space. As musical practice shifted, composers began to specify not only pitch and rhythm, but also loudness by supplying markings for the *dynamics*—the force or power with which the music was to be performed.

Where does timbre appear? The answer is, it does not, or at least, not directly. The intersection of pitch, rhythm, and dynamics will necessarily constrain the range of possible timbres for a given instrument, and thus staff notation *does* reflect timbre indirectly; however, timbre was not used as a dynamic musical parameter in the way that pitch or rhythm were used as dynamic musical parameters, and thus timbre was afforded no place in staff notation, because none was needed, with one notable exception relevant only for vocal music: the text. Other than this, for most of Western history, timbre was exclusively a matter of performance practice.

Of course, I am begging an important question here: what is timbre anyway? While a perfectly legitimate question, this is also a dreadfully complicated one. The modern concept of timbre did not even emerge until the seventeenth century and did not enter common parlance until well into the eighteenth. Rousseau described it in Volume XV of the *Encyclopédie*¹ as the difference between two sounds that can be attributed to neither pitch nor loudness. Even composers who developed sophisticated techniques for the expressive use of timbre were unable to provide a direct definition. In his grand treatise on the topic, rather than saying what timbre *was*, Berlioz was only able to describe its compositional application: a means of *coloring* of the melody, harmony, and rhythm—coloring, but *not* changing.² Another hundred and fifty years of scientific and music-theoretical progress later, the Acoustical Society of America presently defines timbre in much the same way as Rousseau: as “that attribute of auditory sensation which enables a listener to judge that two nonidentical sounds, similarly presented and having the same loudness and pitch, are dissimilar.”

¹“As for the difference which is found between sounds by the quality of timbre, it is evident that it results neither from the degree of lowness [pitch], nor even from that of loudness. . . . for the quality of timbre depends neither on the number of vibrations which make the degree of high or low, nor on the largeness or the force of these same vibrations which make the degree of loud or soft. It will be necessary therefore to find in the corps sonores a third modification different from these two to explain this last property; a project which doesn’t seem to me too easy a thing.” as quoted in Fales (2005, 4).

²“The employment of these various sonorous elements [musical instruments], . . . either for *colouring* the melody, harmony, and rhythm, or for producing peculiar impressions . . . independently of all aid from the three other great musical powers,—constitutes *the art of instrumentation*.” Berlioz (1882), Mary Cowden Clarke’s translation; emphases in original.

1.2 Timbre and Harmony

Well, what do we know? We know that timbre has something to do with overtones. What exactly this means is difficult to say, but it does raise an interesting question: what is the relationship between timbre and harmony? Consider the harmonic partials of an ideal vibrating string.³ Each of those pitches is literally present in the sound, and if you were to isolate and listen to the first five successively, you would be right to call them the members of a C-major triad. Yet when the harmonic partials of the *corps sonore* sound together, we perceive only a single tone having a particular timbre. Where, then, lies the line between timbre and harmony?



Figure 1: harmonic partials (image from Helmholtz (1885, 22))

To answer this, let us break down the listening process itself. Imagine a violinist drawing her bow across her instrument. This action sets in motion a standing vibration in the string, transmitted through the bridge to the resonating body of the instrument and thence into the air, through which the vibrations travel as propagating waves—time-varying atmospheric pressure.⁴ This is what we call *acoustics*. When the acoustic signal reaches the listener’s ear, the ear performs on the signal a sort of biological Fourier transform, thus converting a single acoustic signal into a bundle of neurological ones that we may call discrete auditory sensations—loudness, rhythm, pitch (pitch as in, individual partials), roughness, etc. This process is commonly called *psychoacoustics*.

The ear’s job now complete, it remains up to the brain to construct from these discrete auditory sensations a coherent auditory image in a process Alfred Bregman (1990) calls “auditory scene analysis”. The brain fuses the various partials together into one or more “auditory streams”. These auditory streams are perceptual Gestalts: the harmonic partials of the *corps sonore* fuse into a single perceived sound possessing a particular timbre, loudness, duration, and pitch.

³see Fig. 1

⁴see Fig. 2

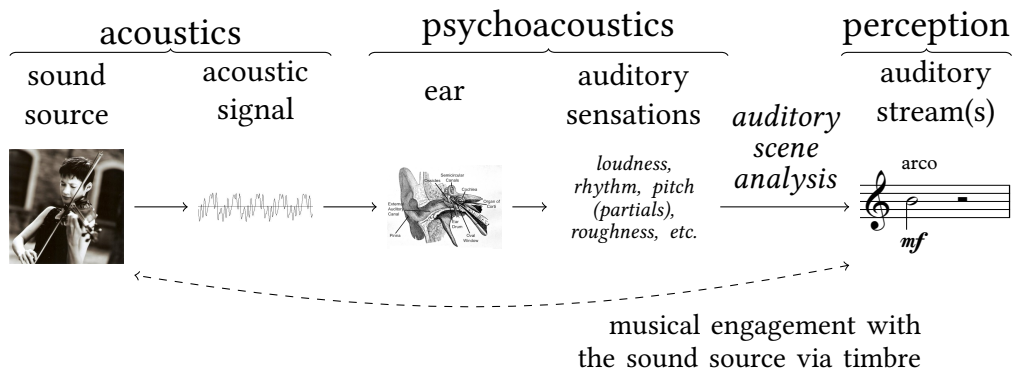


Figure 2: a model of listening

The brain does all of this heuristically by fitting the evidence at hand to known schemata. Or, less technically, the brain takes the available data and from them makes a best guess, and where the data conform more or less to familiar patterns, the brain is all too happy to fill in missing details and smooth out rough edges, a processes ethnomusicologist Cornelia Fales (2002) calls “perceptualization”.

It is here, I would argue, where sensations are grouped into perceived auditory streams, that we find the line between timbre and harmony. Timbre is a property belonging to a single auditory stream; harmony requires multiple streams. But lest you think that I’ve solved it, bear in mind that this line is a terribly fickle line, and in both theory and practice, it is far less obvious than we might like. Alfred Cramer (2002) argues that Schoenberg’s *Klangfarbenmelodie* played upon this line. Rebecca Leydon (2012) shows how identical passages of Crumb can fall to one side or the other depending on the performance. In the same city where Rameau published his *Génération harmonique*, Boulez built IRCAM, and timbre and harmony collapsed into a single, inseparable technique—the line between timbre and harmony dissolved altogether into meaninglessness, replaced by pure spectralist affect.

Thus, I have solved nothing, proposed a line only to blur it. Ask me not what timbre is, for I cannot tell you. And while over the past several years, more and more music theorists have become aware that this is a largely unanswered question, that timbre has been under-theorized and remains greatly in want of the rigorous taxonomical codification that will bring to timbre the same lucidity we now enjoy with set classes and sonata forms, sadly, this has not yet come to pass, and while I am sure there are many fine doctoral students working on

exactly this as we speak, I, unfortunately, am not one of them. So instead, like Berlioz, allow me to speak simply of how timbre is *used*.

2 Timbre, Subjectively

It is through timbre that we perceive our auditory world. That is to say, the timbre of a sound reveals to us how that sound was produced. We hear the tone of a piano, and we know that it was a piano that produced that tone. We hear the roar of a diesel engine and the blare of a horn, and we step back onto the curb, not because we fear the *sound* of an oncoming truck, but because we wish to avoid the truck itself. We hear the voice of our mother saying, “Come here, my child,” and thus we interact with our mother. These engagements are *mediated* through timbre, but they are not engagements *with* timbre. As philosopher Stephen Davies writes,

it is not merely the accident of association that leads us to think of sounds in terms of their sources. Indeed, wider consideration of the evolutionary function of our senses suggests that, rather than being concerned primarily with what impinges directly on the ears, and then with what causes this only by association, their adaptive function is to provide knowledge of those sources. ... Accordingly, we generally experience not the qualities of the sound we are sensing but the physical properties of its source. In other words, we register changes in the sound as changes in the cause of the sound. For example, if a percussionist strikes his instrument with decreasing force, rather than following the subtle, complex acoustic changes that occur, we hear the change in the force with which the instrument is struck.⁵

(For those of you more swayed by quantitative science than the word of a philosopher, a meta-analysis by Giordano and McAdams (2010) corroborates this statement.⁶)

The perception of timbre emerges along with the formation of auditory streams, as raw sensations are molded into perceptions. This is the point where the

⁵Davies (2008, 370–71)

⁶“Overall, the analyses reported here are consistent with current theories of timbre, according to which this multidimensional attribute of auditory sensation is indeed the perceptual correlate of the mechanics of the sound source.” Giordano and McAdams (2010, 165)

higher-order processing broadly known as “cognition” begins, but this is also the first point at which a cycle can be completed, the first point at which the subject perceives the sound *source*, the first point at which the Listener engages with the Listened-to. It is at the moment of perception that this link, this *communication*, is established, and *music* can only begin *after* the signal has passed *through* the pre-conscious process of auditory scene analysis, a process that is heuristic, messy, deeply human, all too human.

As I mentioned before, staff notation represents timbre directly in only a single dimension: lyrics. And this is appropriate—after all, language *is* fundamentally timbral in nature. When we hear language, we hear timbrally, synthetically: individual phonetic elements are combined into linguistically meaningful phonemes. Rebecca Leydon (2012), building upon Bregman (1990) and Fales (2002), provides a powerful demonstration of how the brain maps the sensory input onto known schemata, and thus how we perceive a Gestalt. This is the same basic process that maps many harmonic partials onto the schema of a single complex tone.

Put another way, to hear timbrally is to hear linguistically, or more specifically, to hear a language when we *speak* that language, when we are participants in that game. If we are outsiders to the language, mere spectators to the game, we will instead be bombarded by a discordant cacophony of allophones, a barrage of phonetic shrapnel, the lonely etic perspective. The *emic* perspective is the synthetic perspective, the schema-based hearing, the one that connects the perception with the source.⁷ To not hear the timbre of a sound is to not hear the body that created it. To hear etic harmony where there should only be emic timbre is to miss the music altogether.

And here is the catch: whether or not these individual components are perceptually fused into a coherent Gestalt is not only a question of those elements themselves. There must, of course, also be an awareness of the schema, but also, environmental context matters: the more distractions present, the more difficult it is to attend to the components singly, the more readily the perceptual fusion is triggered. Also, the listening subject’s *will* matters. Play a note on the piano—it is not difficult to pick out the second or third partials if you listen attentively and wish to hear them out; neither is it difficult to hear parallel octaves as a single musical object moving through pitch space, if we allow ourselves to hear and feel it thus. Or, in sum, the premise with which I started: the line that separates the perception of pitch from the perception of timbre is contextual, subjective, and

⁷see Fig. 2

fluid.

3 Analysis

Staff notation lives a double life. Principally, it provides the performers with instructions telling them what they should *do*, and it is their actions that lead to the perceived auditory streams from which the music is made.⁸ The most convenient way for musicians to visually represent these perceived auditory streams is again staff notation. This is what it means to “read music”: to see a score and to mentally “hear” the music without the mediating steps of literal performers or literal ears. This is what we theorists and analysts do: we have learned to silently experience the music by reading the score in a very similar way to how a listener can sonically experience the music by listening to a recording through a pair of headphones.

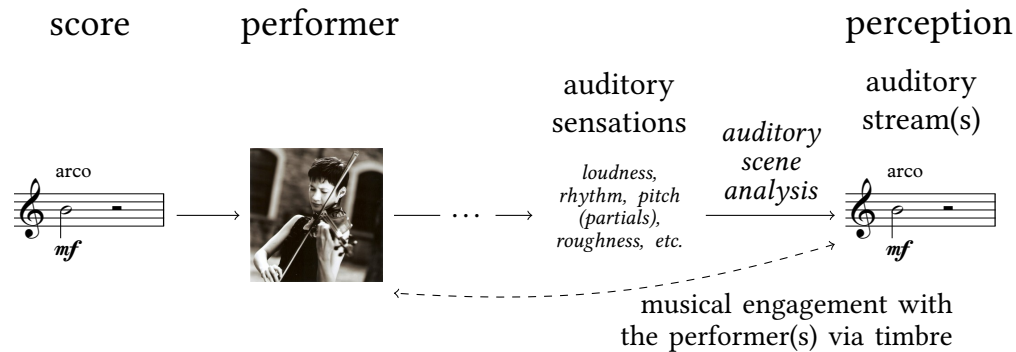


Figure 3: a model of listening (simplified), with score

Or at least, this is what we *should* do. What we generally *do* do is to simply assume that the score *already* represents the experienced music. This is the fundamental assumption made in most analyses—we wish to analyze some piece of music, we assume that the score *is* a representation of that music, and therefore we proceed analyze the score.⁹ But when our analyses valorize the scores *as* the music in this way, they ignore any significance auditory scene analysis may play

⁸See Fig. 3

⁹See Fig. 4

in the process of the score *becoming* music—becoming music through the actions of the performers and the equally active participation of the listeners.

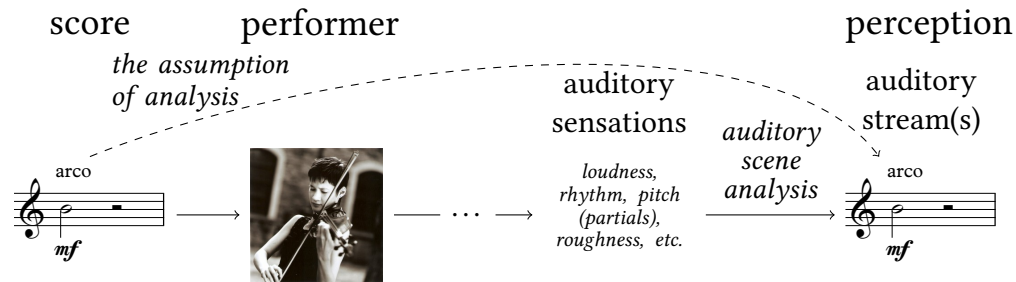


Figure 4: the assumption of analysis

Thus, the analytical neglect of timbre is symptomatic of the assumption of analysis, an assumption that effectively bypasses the role not only of performer, but also—and here is the great irony—of the listener. The timbre of a sound is how that sound came to be, how it came to be *heard*, and it is in that anfractuous coming-to-be wherein music lies. The analytical processes upon which we theorists so depend can lead us to see only the score, but to miss the *music*.

4 The Journey

In my remaining five minutes, let me turn to the symphony I promised to discuss in my proposal. Einojuhani Rautavaara’s symphonies span the second half of the twentieth century,¹⁰ and their stylistic diversity and transformation over time parallel the shifting trends and tastes in compositional practice. They are a microcosm of orchestral music since the second world war, and his eighth can be taken as representative of the neo-romanticism that was so prominent at the close of the last century and which remains a dominant expressive mode today. While Rautavaara *has* composed several more large-scale multi-movement orchestral works over the past 15 years,¹¹ most prudently, “The Journey” is the last one he has labeled a “symphony”. Consequently, by cautiously stopping at eight, Rautavaara is still alive.

¹⁰Symphony No. 1 (1956); Symphony No. 2 (1957/1984); Symphony No. 3 (1961); Symphony No. 4: *Arabescata* (1962); Symphony No. 5 (1985–1986); Symphony No. 6: *Vincentiana* (1992); Symphony No. 7: *Angel of Light* (1994); Symphony No. 8: *The Journey* (1999)

¹¹e.g., *Book of Visions* (2004); *Before the Icons* (2005); *A Tapestry of Life* (2007)

Throughout the first movement, the *divisi* violins frequently move in parallel motion, spanning an octave. The interior space of the octave is filled in with two or three other pitches, depending on the passage. Consider the passage beginning with bar 21.¹² The interior intervals above the bottom note of the octave are a diatonic third, a perfect fourth, and a perfect fifth.¹³

The heavy tonal allusions and triadic harmonies are typical of Rautavaara's late style, as they have been typical of compositional practice for the past thirty years. As I hear it, in this passage, the range, voicing, and parallel motion of the violins weaken, but do not supplant, the perception of the *chordal* planing—each of the five pitches performed by the violins retains small degree of individuality as a perceptual stream.

The image shows a page of a musical score for Rautavaara's Symphony no. 8, first movement, bars 21-25. The score is in 3/4 time with a tempo marking of quarter note = c. 50. It features multiple staves for woodwinds (Ob., Eng. Hn., B. Cl., Bsns., Cbsn., Hns.), strings (Vln. I, Vln. II, Vla., Vc., Db.), and trumpets (Tpts.). The woodwinds and strings play chords, while the trumpets play a melodic line. Dynamics range from f to mp.

Figure 5: Rautavaara, Symphony no. 8, mvt. i, bars 21–25 (some rhythms reduced for clarity)

Consider now a second passage just a few bars later,¹⁴ again, *divisi* violins spanning an octave. (Woodwinds join the strings from bars 33–38, then drop out again as the phrase concludes.) Attempting to tease out the harmonic logic of the passage is tricky. The accompanying voices offer a gentle succession of seventh chords, but over this hangs an inescapable dissonance in the violins—biting major sevenths and minor seconds, an [015] trichord that remains consistent

¹²See Fig. 5

¹³Listen here: <https://youtu.be/XoS7x6OoNw0?t=2m1s>

¹⁴See Fig. 6

throughout the passage. The “journey” here is one of sharp, dissonant structures moving against the accompanying seventh chords, two disconnected strata—the melodic planing of the top voices and the harmonic succession of the bottom voices—ignoring each other, intersecting and overlapping at moments, but far from integrated into a unified harmonic context, and undermining any sense of tonal resolution.

Thus I fall headlong into the assumption of analysis—an ocularcentric dissection of the score that has not bothered to consider the music. Could the analysis I just provided withstand the simple experience of listening to the passage?¹⁵

To my hearing, with the arrival upon bar 28, the four pitches performed by the violins are fairly distinct and easily parsed into four auditory streams. But by bar 29, they perceptually fuse in into a single stream that functions as a *single* pitch with an unusual inharmonic timbre. The winds, with their entrance in bar 33, tease apart this fusion, or at least, they offer the listener that opportunity for a moment; they drop out in bar 38, and their critical support suddenly removed, the four “pitches” fuse quickly into one, and the melody comes to rest upon the root of a simple A \flat -minor triad—no more, and no less.

“But in bar 39, what of the violins’ E \flat and G so clearly marked in the score?” the score-analyst might cry. “Are you claiming these pitches do not matter?” No. I am not claiming that these pitches do not matter. Rather, I am claiming that these *are not pitches*. The E \flat and the G that are printed in the score are only there because that was what Rautavaara needed to tell some of the violinists to do so that the A \flat would have the right color, the right timbre.¹⁶

Music is something that you *hear*, not something that you see; or, as Joel Krueger (2009) argues, “not simply constellations of acoustic properties or ‘pre-ordained gestures’ . . . collectively transferred from composer to listener.”¹⁷ Rather, music is something *enacted*, or as Davies puts it, music is “the auditory bodying forth of human action.”¹⁸ There is neither E \flat nor G there; there is only the A \flat , entimbred—entimbred at the request of the composer by the actions of the performers. Timbral listening is an engaged *participatory* listening—contextual, subjective, fluid, and to a surprising degree, willful. To not hear timbre is to not engage with the embodied musicians who perform it, which is to say, to miss the music altogether.

¹⁵<https://youtu.be/XoS7x6OoNw0?t=2m37s>

¹⁶See Fig. 7

¹⁷Krueger (2009, 120)

¹⁸Davies (2008, 369)

27 ($\text{♩} = \text{c. } 50$) 28

Vln. I
Vln. II
Vla.
Vc.
Db./B.Cl./
Cbsn.

33

Tbns.
Hns.
Vln. I
Vln. II
Vla.
Vc.
Db./B.Cl./
Cbsn.

37

Tbns.
Hns.
Vln. I
Vln. II
Vla.
Vc.
Db./B.Cl./
Cbsn.

Figure 6: Rautavaara, Symphony no. 8, mvt. i, bars 27–39 (reduced for clarity)

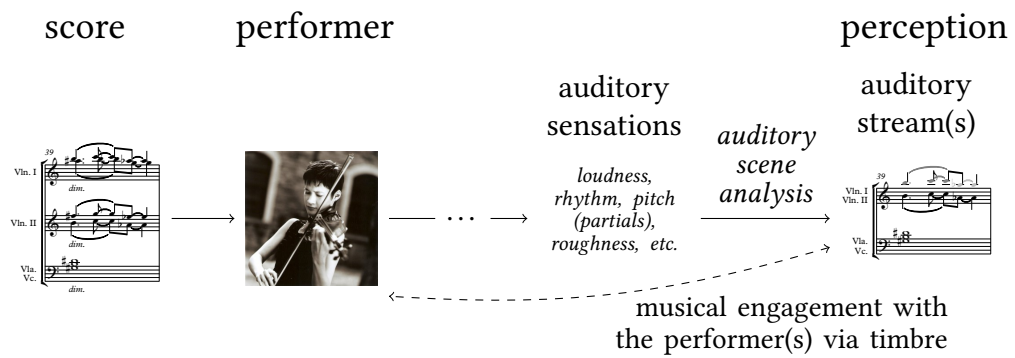


Figure 7: Listening to Rautavaara

References

- Berlioz, H. (1882). *A Treatise on Modern Instrumentation and Orchestration*. London: Novello, Ewer and Co.
- Bregman, A. S. (1990). *Auditory Scene Analysis: the Perceptual Organization of Sound*. Cambridge, Massachusetts: The MIT Press.
- Christensen, T. (1987, Spring). Eighteenth-century science and the “Corps Sonore:” the scientific background to Rameau’s principle of harmony. *Journal of Music Theory* 31(1), 23–50.
- Cramer, A. (2002, Spring). Schoenberg’s *Klangfarbenmelodie*: A principle of early atonal harmony. *Music Theory Spectrum* 24(1), 1–34.
- Davies, S. (2008, October). Musical works and orchestral colour. *British Journal of Aesthetics* 48(4), 363–375.
- Fales, C. (2002, Winter). The paradox of timbre. *Ethnomusicology* 46(1), 56–95.
- Fales, C. (2005). Listening to timbre during the French enlightenment. In *Proceedings of the Conference on Interdisciplinary Musicology (CIM05)*, pp. 1–11.
- Fastl, H. and E. Zwicker (2007). *Psychoacoustics: Facts and Models*. Number 22 in Springer Series in Information Sciences. Berlin: Springer-Verlag.
- Giordano, B. L. and S. McAdams (2010, December). Sound source mechanics and musical timbre perception: Evidence from previous studies. *Music Perception: An Interdisciplinary Journal* 28(2), 155–168.
- Helmholtz, H. (1885). *On the Sensations of Tone as a Physiological Basis for the Theory of Music*. translated by Alexander J. Ellis. London: Longmans, Green, and Co.
- Huron, D. (1991, Winter). Tonal consonance versus tonal fusion in polyphonic sonorities. *Music Perception* 9(2), 135–154.
- Krueger, J. W. (2009). Enacting musical experience. *Journal of Consciousness Studies* 16(2–3), 98–123.
- Leydon, R. (2012, June). Clean as a whistle: Timbral trajectories and the modern musical sublime. *Music Theory Online* 18(2).