

# Consciousness and Mental Qualities for Auditory Sensations

Adriana Renero

Philosophy, The Graduate Center, City University of New York

[arenero@gc.cuny.edu](mailto:arenero@gc.cuny.edu)

## Introduction and Preliminary Considerations

How are we conscious of our auditory sensations and perceptions?

Recent philosophical theories of sound and auditory perception tend to focus on various aspects of sound itself; e.g. (1) the ontology of sounds, or whether sounds are secondary qualities, waves, or some type of event, (2) the mere content of auditory experience, or what sorts of properties we perceive and how sounds are individuated in virtue of their temporal patterns and spatial locations, and (3) the sources and specific location of sounds.<sup>1</sup> But these theories are limited only to questions regarding how we become aware of the physical sound conceived, for example, as a property of an object (Pasnau, 1999), an object-like (Nudds, 2001), an event-like individual (Casati and Dokic, 1994; O'Callaghan, 2007, 2009), a pure event detached from its physical causes (Scruton, 1997), or as a stable property (Kulvicki, 2008).

Although the contribution of these accounts has been extremely significant for the development of a philosophy of sounds and audition, none tackle the specific question of how our consciousness of auditory states arises. Here I propose an account of auditory consciousness, which includes both an account of the qualitative character of auditory states and how those states become conscious. This account aligns with extant theories (a kind of perceptual-role theory and a metacognitive theory of consciousness) which have been independently motivated, and also extends those theories by advancing proposals specific to this neglected but important domain.

One motivation for this inquiry is that auditory states have a puzzling nature, which sets audition apart from other modalities. As opposed to ordinary tangible, static, or material objects which are generally grasped through vision or touch, auditory states have to do with particulars that occur or change along with time, and thus *qua* auditory states they are constitutively temporal, or exhibit temporality when those extended particulars or phenomena such as change, succession, and persistence occur.

Auditory states also entail momentary experiences, variations, anticipations and predictions, and comprise retention of recent past events. Yet even though auditory states might seem at first sight to be too elusive and complex to be examined, they are nevertheless significant for the study of consciousness. Specifically, being conscious of our auditory states as opposed to being conscious of a merely particular, object, or auditory feature is necessary for the study of our mental lives—it is meaningful not only

---

<sup>1</sup> For a discussion on the ontology and sources of sounds: Pasnau (1999); Casati and Dokic (1994); Dokic (2007); Hamilton (2009). I overlook the problem of auditory hallucinations, auditory illusions, and any reference to 'auditory objects' in recent definitions of sounds: O'Callaghan (2008); O'Callaghan and Nudds (2009).

for our subjective experiences *per se*, but also for what we process while thinking about those very experiences.<sup>2</sup>

The prevalent theories of consciousness, on the other hand, tend not to discuss in detail the auditory modality involved in our experiences. In contrast to the vast research regarding visual states, auditory states had been constantly neglected in the philosophical literature on consciousness.

One might inquire what kind of auditory particulars is being engaged in the auditory experience that I investigate here. Unquestionably, there is a difference between our having an auditory experience of complex sounds vs. our having such an experience of less complex or simple sounds. Although different audible features such as pitch, timbre, loudness, and spatial location can be similarly processed by listening to a speech, a lions' growling, a horn, a yell, an automobile tires' screeching, clicks, percussive noises, or a musical piece, there is a peculiar distinction between hearing at this time the sound of a yell or a simple click, vs. hearing through time—and at the same time—changes, successions, variations, and repetitions of musical sounds. Given that musical sounds are inherently defined as complex relationships determined by length and width—thus discerning and defining them are challenging enough—, I concentrate on analyzing the auditory experience of them.

My aim is to explain how audition goes from being not conscious to being conscious; in other words, how our consciousness of auditory states arises. That is, what is it for a subject to be conscious of her auditory states with particular qualitative distinctions? In searching for an answer to this question, I examine a range of auditory mental phenomena (pitch, timbre, loudness, and spatial location) with their particular properties. I argue that we are able to capture qualitative distinctions of our auditory sensations and to be conscious of our auditory states with specific distinctions while having an auditory experience.

To develop an approach to auditory experience, it is helpful to build on approaches to consciousness that have already been independently motivated in the literature. Here, I draw on David Rosenthal's high-order-thought (HOT) theory of consciousness plus his quality-space theory (QST) of mental qualities (2005; 2010).<sup>3</sup> I will argue that these theories lend themselves to audition, and their successful application there, provides further support for their explanatory adequacy. QST is intended to explain how we discriminate among perceptible properties that figure in qualitative experience by locating those properties in a quality space. The HOT theory offers an explanation of how we are aware of conscious qualitative states.<sup>4</sup>

---

<sup>2</sup> I overlook the debate on 'inverted spectrum'—i.e. whether two subjects being physically alike are different in terms of what it is like for them to have certain conscious states—and 'zombies'—i.e., whether these creatures have conscious states or not.

<sup>3</sup> Mental qualities are 'the properties of perceptions in virtue of which an individual can discriminate among the relevant perceptible properties' (Rosenthal, 2010: 378). Rosenthal does not call mental qualities 'qualia', because on his view they are not necessarily conscious—they might be conscious and they might not be. However, qualia are defined as 'properties of conscious experiences [or states] in virtue of which there is something it is like to have experiences' (Mandik, 2010: 97) or to have a distinctive 'feel' (Nagel, 1974).

<sup>4</sup> Qualitative states 'are those mental states in virtue of which we respond mentally to the properties of environmental objects and processes' (Rosenthal, 2005: 177). Here I address sensations and perceptions. 'A sensation is a 'preconceptual conscious state that figures among the raw materials out of which a perception may be formed... (Mandik 2010: 108–9).

This project takes advantage of the particular merits of both the HOT theory of consciousness and the QST of mental qualities. I show that the HOT theory, supplemented by the QST, can be profitably employed for the auditory modality. Specifically with respect to the qualitative character of sounds, and to the way we understand the consciousness of auditory sensations with their respective qualitative distinctions.

No one has shown the applicability of these theories in tandem to the auditory modality. Rosenthal does not analyze how specifically auditory states with their auditory qualities work within his account, and proponents of HOT and QST have not discussed the auditory case in depth. Moreover, specific details about how this domain works cannot be trivially inferred from HOT or QST. Here, I will offer specific proposals about the qualities of pitch, timbre, loudness, and spatial location for complex sounds. These extend well beyond what is presented in work on HOT and QST, and some of my proposals could be adapted to fit with other theories of consciousness as well.

The consequence of making use of Rosenthal's theories in this respect is, not only the expansion of Rosenthal's theories (as I show at §§2-3)—, but also a supplement to existing philosophical theories of auditory perception; in particular, to the individuation of auditory experiences according to the way we perceive audible features, and the way we create quality-spaces for our auditory states.

In order to analyze how our consciousness of auditory states arises, my strategy is to use QST to show how we can discriminate among an array of similarities and differences of mental qualities such as pitch, timbre, loudness, and spatial location. I further take HOT theory to explain what it's like to be in auditory states with those particular qualitative distinctions. After discussing how we become conscious of auditory states with those qualities, I show how acquiring experience and conceptual resources for the qualities of our auditory sensations might affect the way we become conscious of auditory sensations with fine-grained qualities. In order to show the supplemented theory at work, I close by observing how a hypothetical experienced listener in possession of a refined musical conceptual vocabulary may describe and report her higher-order-thoughts about her musical experience.

The paper is divided into three sections: I begin by pointing out some of the basics of HOT and QS theories (§1). After that, I will argue for my own proposals to show how the QST of mental qualities works for the musical subset of the auditory modality (§2). Finally, I will show how one's higher-order-thoughts capture qualitative distinctions of auditory sensations, and what factors might contribute to one's consciousness of auditory states while having a musical experience (§3).

## **1. Higher-Order-Thoughts and Quality-Space**

HOT theory maintains that qualitative states sometimes occur without their being conscious, as is the case of subliminal perception—i.e., when a stimulus is undetected or just perceived below the threshold of consciousness, but the stimulus still implies a sensory response (Rosenthal, 2005: 177–179).

HOT theory further claims that although we might be conscious of something (a merely particular, object, or event) by sensing it—i.e. by having first-order states<sup>5</sup>—, we do not become conscious of our qualitative state just by sensing. We require a further thought that makes us aware of being in that sensing state—a higher-order state. So, having a higher-order-thought is the way of being conscious of ourselves as being in that state.<sup>6</sup> This is a combination of *transitive consciousness* or being conscious of something as being present and *state consciousness* or that state itself being conscious. Thus, a state is conscious if one is conscious of that state in some suitable way—that is, by having a thought about oneself having the experience in question (Rosenthal, 1997: 741; 2004: 18–22; 2005: 182–5).<sup>7</sup>

So, *what it is like* for us to be in a distinctive qualitative state depends on how we are conscious of that state. However, in order to explain conscious of being in a qualitative state, and to describe that state in terms of corresponding qualities, a kind of perceptual-role theory that accounts for the mental qualities that that state exhibits is necessary. Hence, I appeal to Quality-Space Theory.

QST discusses the qualitative character of our mental states and accounts for a way in which we are able to discriminate among an array of mental qualities for any sensory modality (Clark, 1993; 2000; Rosenthal, 2005; 2010). Indeed, QST gives an account of mental qualities that is independent of those qualities we can access by consciousness (Rosenthal 2010: 373–76).

In order to account for mental qualities *per se*, this theory builds on the quality-space of discriminable perceptible properties—rather than on physical properties of the stimuli. So the distinction between physical properties or characteristics accessible by perception and the corresponding mental qualities is relevant.<sup>8</sup> That is, in the first instance, we discern similarities and differences among the perceptible properties of any sensory modality, and thus we construct a quality-space of those properties based on ‘just noticeable differences’ and in respect of their location, position, or on ‘how distant each property is from every other’ property (Rosenthal, 2005: 198–222; 2010:

<sup>5</sup> Rosenthal argues against first-order theories such as Dretske’s (1997), which posit that a state is conscious if it simply makes one conscious of something; such theories arguably cannot explain non-conscious mental states.

<sup>6</sup> At this point the theories of Carruthers (2000) and Rosenthal (1997; 2005) differ. While Carruthers stresses that in order for a mental state to be conscious, it is necessary to have a disposition to be the object of such a higher-order representation, Rosenthal stresses that the importance of the higher-order representation to occur involves actually being the target of a higher-order representation.

<sup>7</sup> For criticism to HOT: (a) *the targetless problem*: Byrne, 1997: 103–29; Levine, 2001; Block, 2009; 2011; (b) *the cognitive overload problem*: Carruthers, 2000: 221–22; (c) *the richness problem*: Block, 2011. For responses to criticism: (a) the dental-fear case: Rosenthal, 2005: 127, 172, 209–11; 2010; 2011; Weisberg, 2011; (b) the cocktail-party effect: Rosenthal, 2005:111 fn. 10:128; at §3 I show that auditory higher-order-thoughts are not isolated but rather come in clusters; (c) at §3 I show that HOT theory in tandem with QST allows for highly detailed discriminations of our auditory experiences by use of comparative thoughts. For empirical support of HOT theories: Lau and Rosenthal (2011), Lau and Brown (forthcoming), Brown (2012). For reasons of space I omit this discussion.

<sup>8</sup> For a discussion on this distinction see Rosenthal, 2010: 379. For specific distinctions between auditory-property and auditory-quality see my §2. For a discussion on third-person/first-person perception, ‘quality inversion’, ‘zombies’, ‘explanatory gap’, and ‘hard problem’: Rosenthal 2010: 368–72; 380–81; my fn.2.

377–80).<sup>9</sup> Derivatively from that process, the quality-space ascertains the discrimination of the corresponding mental qualities.<sup>10</sup>

It is worthwhile emphasizing that ‘location’ according to the quality-space theory does not refer to the spatial location of the physical properties or characteristics; it refers only to the properties with respect to the way we distinguish among them.

## 2. Auditory Quality-Space

With these two preliminary views explained, I can now fill in the details of how we can create an auditory quality-space. In this section, I argue that QST lends itself to audition, and its successful application here, provides further support for its explanatory adequacy. I show that QST offers a useful platform for the qualitative character of sounds, and in capturing the scope of fine-grained differences among them. I further show that QST allow us to map similarities and differences among, in the first instance, noticeable audible properties, (also called) acoustic or physical properties, such as the rate of sound vibration, the quality of sound that distinguishes one instrument from another, and the level of strength or energy of sound vibrations (in simple terms these can be identified, respectively, as frequency, tone-color, and amplitude); and derivatively from that mapping, the corresponding mental qualities of pitch, timbre, and loudness.<sup>11</sup>

It is important to see that sounds have at least frequency, tone-color, and amplitude independent of being perceived. For simplicity’s sake, I will write the audible *physical property* of the rate of sound vibration (or frequency) as ‘pitch<sub>p</sub>’, and the perceived characteristic or its related *mental quality* as ‘pitch<sub>m</sub>’—i.e. the quality of “highness” or “lowness” of sound that we perceive—, and do the same for timbre and loudness. Notice how, in my discussion, I keep the distinction between physical properties and mental qualities.

First, if we accept that QST can be employed profitably for the auditory modality, we might say that auditory mental qualities are the properties of auditory sensations in virtue of which we are able to discern among the range of audible properties of sounds. QST would allow us to map similarities and differences among, in the first instance,

<sup>9</sup> According to Rosenthal (2010: 378), ‘the distance between any two perceptible properties is a function of how many properties between the two the creature can discriminate’.

<sup>10</sup> For objections to QST see *the experiential holism problem*—e.g., ganzfeld, diachromatic and color agnostic cases: Prinz 2012: 131–132. QST advocates might response: since the core of our discriminative ability is to grasp perceptual *differences* among *distinct* types of stimulus, and since ganzfeld is a phenomenon of perceptual deprivation of the stimulus, and no comparative component is at stake in diachromatic and color agnostic cases, the previous objection does not mean a serious challenge for QST. For reasons of space I omit this discussion.

<sup>11</sup> Pitch, timbre, and loudness are identified as audible *features* in the philosophical literature on sounds (O’Callaghan, 2008). In the literature on music, these are called *properties* of sounds (Randel, 1986; Kerman, 2008). I use ‘audible properties’ as equivalent to ‘perceptible properties’, in Rosenthal’s terms. I reserve ‘quality’ for ‘mental quality’, ‘auditory mental quality’ or the psychological dimension. Neither the use of ‘property’ nor ‘quality’ here is an attempt to equate them with Locke’s notions of ‘sounds as properties of the bodies’, nor of ‘sensible qualities possessed by bodies’. Nor do I address the controversy among property, wave, and event views (O’Callaghan, 2007; 2009; Kulvicki, 2008).

noticeable distinct ‘pitches<sub>p</sub>’, and derivatively from that map their corresponding ‘pitches<sub>m</sub>’.

How, specifically, would QST account for the qualitative character of sounds and their fine-grained differences? Rosenthal (2005: 198) points out that ‘the ways the various mental qualities resemble and differ from one another are homomorphic to the ways the corresponding perceptible properties of... events resemble and differ. A sensation of the sound of a trumpet, for example, resembles a sensation of the sound of a woodwind more than it does a sensation of the sound of a violin.’

Since he does not elaborate on audition, I infer (A) that Rosenthal conceives of sounds as *events* of a certain kind; and (B) that the resemblance and difference among audible properties is where we should look to individuate the qualitative character of sounds. If (A) is correct, we should take it that sounds are events that occur over time. They are concrete individuals with durations: a beginning, middle, and an end, which can therefore be located in space and time.<sup>12</sup> If (B) is correct, it is necessary to explain *what* exactly we can discriminate of the qualitative character of a sound. We can, at the very least, discriminate between timbres, pitches, and varying degrees of loudness, for example.

To expand: as is well-known, sound events can be classified according to the several properties we can perceive, such as pitch, timbre, and loudness. These are the sorts of properties that allow us to classify sounds as being of a certain type. Our auditory quality-space would have then at least three dimensions: we can discriminate among pitches<sub>m</sub>, timbres<sub>m</sub>, and loudness<sub>m</sub>. It is important to observe that the auditory system is able to detect together and simultaneously different sounds and properties of sounds, and so the brain seems to process several distinct properties of sounds at once. Also, QST would allow us to map pitch, timbre, and loudness in a multidimensional space. However, for simplicity’s sake, I only examine audible properties and corresponding mental qualities *individually*.

In support of this suggestion, let us first explore the rates of sound vibration; in scientific terminology: frequencies, which are measured in cycles per second called Hertz (Hz), or what I call ‘pitches<sub>p</sub>’.

The total range of frequencies that unimpaired ears detect goes from 16 Hz to 20,000 Hz; this range varies according to age and other factors. Interestingly, musicians themselves almost never use the total range of frequencies—like that which we hear in the sliding scale of a siren, for instance, when it starts low and goes higher and higher until it is out of earshot. Instead, musicians use a limited number of fixed pitches that extend from ca. 20 Hz to ca. 5,000 Hz. These pitches are calibrated scientifically, with European-styled orchestras tuning to a pitch with a frequency of 440 cycles per second which they label ‘A’ and then assemble into a scale (Kerman, 2008). The difference, or distance, between any two pitches in such scales is called an *interval*; given its peculiar character, the most important interval is the octave (Randel, 1986).

If this is correct, QST, then, can describe a pitch<sub>p</sub> with respect to how among an array of them we discriminate. We can sense the pitch<sub>p</sub> of an individual note that fixes

---

<sup>12</sup> I borrow the initial conception of sounds considered as *events* from Casati and Dokic (1994), and O’Callaghan (2007; 2009). A feature of these views is that sound events have *distal locations*—i.e., occur in or near their sources. I don’t address the problems of sound’s apparent spatiality, ‘pure event’ nor appreciating music independently from its sources, see Scruton (1997), Hamilton (2009).

its position in the scale depending on how high the note is in respect to another—e.g. one pitch<sub>p</sub> is relatively low at 440 Hz, while it is an octave higher at 880, and it is higher still at 1320 Hz. Evidently, we do not need to know the specific hertz in order to sense those distinctions.

The distance between any two pitches<sub>p</sub> in a quality-space depends on how many similar qualities of pitch<sub>m</sub> we are able to discern. For example, we can discriminate among a variety of pitches<sub>m</sub> that range from low to increased frequency, such as from 440 Hz, then passing through 880, 1320, 1760, 2200, and so forth at multiples of the fundamental frequency<sup>13</sup> until finally on to 5280 Hz, as shown on the x-axis of the figure below.

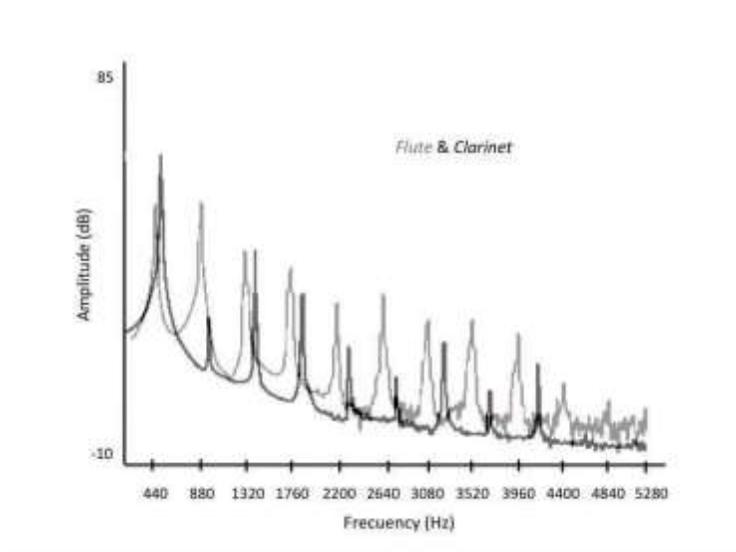


Figure 1, I borrowed two diagrams of Lotto & Holt (2010) and made them converge in order to illustrate a range of discriminable similarities and differences among qualities of two different woodwinds—spectra for a flute and clarinet playing the same note (*ceteris paribus*). We can capture how fast the frequency is repeated, how extensive the wave for the same frequency is, and how the amplitude (timbre) is different depending on the instrument played. We can compare the qualities of each with considerable precision.

Next, how would quality-space map characteristic tones captured by *timbre*—also called tone-color, or the quality of sound that distinguishes one instrument or voice from another? Given a certain pitch and loudness, sounds differ depending on the instruments and voices that generate them. When two instruments play the same note or pitch, we can discern the quality-space between them; for example, between a clarinet-like sound and a flute-like sound (Figure 1).

Differences in timbre like these are clearly heard before a typical concert begins. In order to tune, the oboe plays an ‘A’ and all the instruments follow by playing the same note. However, even though each instrument plays the same note, we can still hear differences in the timbres of every instrument or sometimes by group: strings, woodwinds, brass, and percussion. Even people who do not know the names of the

<sup>13</sup> The fundamental frequency is ‘the lowest frequency or tone in the harmonic series and hence the first harmonic, i.e. the frequency of which all remaining frequencies in the series are integral multiples’ (Randel, 1986: 330).

instruments can distinguish between a smooth, rich sound, like that of the violin, and the hollow, brassy, bright sound of a trumpet. Thus, although some sounds are similar, we are able to sense, say, the timbre<sub>p</sub> of a flute and the timbre<sub>p</sub> of a clarinet; the mental quality of each timbre—i.e. timbre<sub>m</sub>—enables us to make that discrimination.

The specification of frequencies and relative amplitudes are easy to sense and measure when a particular instrument is played. However, this is by no means the only clue used to identify instrumental sounds. Timbre<sub>p</sub> has a peculiar character sometimes difficult to discern—i.e., the characteristic sound results, largely, though not exclusively, from the relative intensities or strengths of the partials or overtones.<sup>14</sup>

Kerman (2008: 4) observes '[p]iano strings... vibrate not only along their total length, but also simultaneously in half-lengths, quarters, eighths, and so on. ...The amplitudes of these fractional vibrations, called partials... [or] overtones..., are much lower than the amplitude of the main vibration. Indeed overtones are not heard as new pitches, but as part of the string's basic or fundamental pitch. The amount and proportion of overtones are what give a sound its characteristic tone-color' or timbre.

When timbres<sub>p</sub> are identified by their location in a quality-space—that is, relative to timbres<sub>m</sub>—this audible property contains in itself overtones<sub>p</sub> in an intimate relation with the fundamental pitch<sub>p</sub>, and also loudness<sub>p</sub>. That is, by sensing timbre<sub>p</sub> our auditory quality-space can map timbre<sub>m</sub> as a singular event from which we could discriminate, albeit with difficulty, its overtones<sub>m</sub>,—e.g. an octave, low and high pitches<sub>m</sub>, and a certain loudness<sub>m</sub>.

In other words, the property of timbre is intrinsically a simultaneity of sounds together—say, a timbre<sub>p</sub> with its respective pitch<sub>p</sub> or frequency, and occurring at a certain loudness<sub>p</sub> or with a specific amplitude distinct of the main. If we were able to discern among overtones<sub>p</sub> without any difficulty, the quality-space we construct of them would be that in which we could experience overtones<sub>m</sub> discerning and capturing the distance between octaves, the differences of its particular or its respective pitches<sub>m</sub>, and its corresponding changes of loudness<sub>m</sub> independently of the main amplitude and of the fundamental pitch<sub>p</sub>.

It is important to clarify that the physical property of timbre involves several amplitudes, one for each octave. If we consider, for example, “each” amplitude in an independent way, this would be an element of a one-dimensional space. But the timbre<sub>p</sub> itself is an element of a multidimensional space because it combines all those amplitudes and frequencies. Since we can distinguish among timbres<sub>m</sub> and presumably not in terms of the frequencies<sub>p</sub> and amplitudes<sub>p</sub> that constitute each timbre<sub>p</sub>, it may indeed make sense to group the frequencies together and regard each distinguishable timbre<sub>m</sub> as a single event describable by location in a multidimensional space.

Finally, our quality-space can also classify the intensity of energy or level of strength of sound vibrations—which in scientific terminology is called amplitude, and it is

---

<sup>14</sup> Overtone is 'a secondary vibration in a sound-producing body, which contributes to the tone-color [or timbre]: also called *partial* (Kerman, 2008: 431). We identify it when running up the white keys on the piano, where a series of successive pitches is sounded; it seems as if the pitch “duplicates” an earlier pitch, but a higher level. This new pitch doesn't sound identical to the old one, but somehow the two sounds are very similar. They blend extremely well; they almost seem to melt into each other. When strings vibrate to produce sound they vibrate in partials or overtones; that is, not only along their full length but also in halves and other fractions. A vibrating string that is half as long as another will reinforce the latter's strongest partial. This reinforcement causes the duplication effect of octaves.

measured in decibels—, or what I call loudness<sub>p</sub> and the class of gradations and changes in loudness<sub>p</sub>. I here use ‘dynamics,’ as musicians do, to refer to the level of sound, or amplitude of sounds and gradation of loudness.

Our quality-space likewise map dynamics<sub>m</sub> and fine-grained gradations in dynamics<sub>m</sub> that go from *pianissimo* or very soft—passing through *piano*, *mezzo piano*, *mezzo forte*, and *forte*—, to *fortissimo* or very loud. Notice that it can also map changes in dynamics<sub>m</sub> which can be sudden—or *subito*—or gradual as in a *crescendo* or *diminuendo*—i.e., gradually getting louder, and gradually getting softer. Moreover, aspects of execution such as *staccato* or *legato*—i.e., a note of shortened duration or notes tied together.

Dynamics<sub>m</sub>, or the perceived intensity of a sound vibration, depends upon the sound pressure, the frequency and duration of the sound, the absorption and the reflection of the sound by the air and surrounding objects, and the distance between the listener and the sound source (O’Callaghan and Nudds, 2009).

We have been able to individuate fine-grained auditory mental qualities (pitches<sub>m</sub>, timbres<sub>m</sub>, and loudness<sub>m</sub>) by their location in the relevant quality-space we create. I demonstrate this view in action by commenting on a musical fragment (at §3). As we will see, an auditory state is conscious when we are aware of being in that particular state with its distinctive auditory mental qualities.<sup>15</sup>

## 2.1. Auditory Quality-Space for Sound Location

One might ask whether QST would allow us to identify sounds as spatially located, and thus whether we could map the location of sounds by distance and direction. We have seen that auditory mental qualities are the properties of auditory sensations in virtue of which we are able to discern among a range of audible properties of sounds. We have also seen that QST allows us to map similarities and differences, and, derivatively from that mapping, their corresponding auditory mental qualities. In this section, I claim that we are able to individuate auditory mental qualities relative to their spatial locations by their particular *position* in the relevant quality-space we create. My proposal not only explains the way we discriminate among auditory spatial *locations* (hereafter ‘sound locations’) and the way this feature works in an quality-space, but also expands QST itself to auditory location.

Two preliminary clarifications are necessary here:

(A) It is important in this context not to confuse ‘position’ with ‘location.’ ‘Position,’ according to the QST, does not refer to the location of physical properties or characteristics, the way sounds are located in physical space or the way they emerge from different directions. To the contrary: ‘position’ here refers only to the properties with respect to the way we distinguish among them.

(B) It is important not to confuse (i) ‘space’ as that being used to indicate the (mental) quality-space that we built of auditory properties based on their differences or the distances between them—that is, in virtue of getting their corresponding mental

<sup>15</sup> Remember that there is a distinction between an auditory state with an auditory mental quality *being conscious* and a state without it—since mental qualities can also occur without appearing in the subject’s stream of consciousness.

qualities—and (ii) ‘space’ as that being used by philosophers of sound and auditory perception to designate sounds as located in the (physical) space or to describe our spatial hearing ability. Sounds conceived of as events can be located either where their material sources are (distal), where the listener is (proximal) or somewhere in the intervening space between the listener and the resonating object (medial).<sup>16</sup>

For simplicity’s sake, I will write the *physical property* of the sound location as ‘sound location<sub>p</sub>’, and the perceived characteristic or its related *mental quality* as ‘sound location<sub>m</sub>’. But how does quality-space could allow us to classify the audible aspect or property of an event intrinsically related to its sound source—what I call sound location<sub>p</sub>?

First, auditory scientists and acousticians believe that the auditory system allows us to orientate ourselves and to discern the location of sound sources with certain precision by making use of several cues. These cues include the estimation of distance and velocity, the intensity, timing measurement, reflection, direction, and spectral information of sounds (Blauert, 1997; Schnupp, Nelken, and King, 2012).<sup>17</sup>

Even as we individuate properties of auditory events as being located in the physical space—i.e. sound locations<sub>p</sub>—, we discriminate simultaneously among other properties of those events as well. That is, when a particular sound is produced by a certain source, we are able to estimate the distance from it to its source by sensing other properties, such as the loudness and the sound spectrum—e.g. distant sound sources have softer loudness or deaden sounds than closer sources; high-frequencies are more rapidly absorbed by the air than low-frequencies.

Other factors may contribute to locate the sound source with certain accuracy: head movement can improve the accuracy of location stereo hearing. When the source is directly located in front or behind the listener, accuracy varies 1° or less; but if the source is located to her left or right, it’s around 15° (Pitt, 2010).

We are also able to estimate other characteristics of sounds thanks to the source’s movement or the listener’s movement. For example, when the listener moves from side-to-side between two static sounds sources, the auditory events in the distance appear “to move” slower than those close to the listener. The number of sound sources, voices or instruments playing the same note or different notes simultaneously may also influence the way we detect sound events; if two separate speakers play different notes we are able to distinguish the two sounds. If they produce the same note, it is typical that we hear a single sound. Yet if a single speaker produces different notes simultaneously, we typically hear two distinct sounds (Blauert, 1997; O’Callaghan, 2008).

Aurally experiencing different sounds as being located in a spatial field is given as a function of their different physical spatial locations—i.e. where the locations of their sound sources are—, and in how the listener is able to make the corresponding

---

<sup>16</sup> In agreement with Casati and Dokic (2009: 109–10), O’Callaghan (2009: 28), and consistent with our phenomenology, distal theories seem to be correct. However, two distal views are at stake: the *relational* event theory, which claims that sounds are events involving both the source and the surrounding medium (O’Callaghan), and the *located* event theory, which claims that sounds are located at their sources—i.e., since sounds supervene on the physical processes in them, sounds are monadic events happening to material objects (Casati and Dokic). For reasons of space I omit discussion on these theories.

<sup>17</sup> For the neural processing that underlies spatial hearing see Schnupp et.al. (2012).

distinctions in terms of distance and direction. In order to make those distinctions, the listener can move her head, focus on loudness, and compare the properties of the auditory events according to their “position” in the relevant quality-space that she builds.

In other words, we can sense sound locations<sub>p</sub> relative to their “location” in the physical space depending on their noticeable distances (how far apart sound sources are) and on their input directions (vertical, horizontal, central, or lateral). We can then discern sound locations<sub>m</sub> with respect to what among an array of sound locations<sub>p</sub> we are able to discriminate.

In terms of directions, we are able to sense vertical and horizontal sound locations<sub>p</sub>, or lateral and central sound locations<sub>p</sub>, and the mental qualities of each one enables us to make the corresponding discriminations. So our quality-space can map those aurally perceived properties—sound locations<sub>m</sub>—as going from *above* to *below* (or the reverse), from the *right* to the *left* (or the reverse), from the *front* to the *back* (or the reverse), from *side-to-side* diagonally, in rotations, etc. —and can thereby discern among those differences.

In terms of distance or spatial arrangement (how “close to” or “far away from” their sources auditory events are), our quality-space can also map the perceived characteristics or their related mental qualities. To put it in simple terms, the position or difference between any two sounds locations<sub>p</sub> in a quality-space depends on what different qualities of sounds locations<sub>m</sub> we are able to discern —e.g. whether one is *close* or *far*. Notice that singling out and referring to a specific location by a demonstrative (*that* sound..., *those* sounds...), may suffice when we listen to a sound coming from a particular distance or direction and attend to the sensation that results from it.

The relevant point is that since we are able to sense sound locations<sub>p</sub>, the quality-space that we construct is that in which we can experience sound locations<sub>m</sub> depending on a set of sound’s properties and cues. To illustrate: let us allude to Karlheinz Stockhausen’s *Oktophonie* and concentrate only on a few details of the electronic accompaniment to explain how QST may account for sounds locations<sub>m</sub> (relative to directions) of our experience as derived from the sounds spatially located.

Briefly, *Oktophonie* is structurally developed from a serially-derived technique that comprises projection, expansion, and multiplication of sounds. Properties such as pitches, timbres (which are created electronically), and dynamics are succinctly joined to the melodic succession which defines both the small and large-scale architecture of the piece. In general terms, it is constituted as a cube in an auditorium of eight loudspeakers or sound layers (I-VIII) located in each vertex, with four loudspeakers placed at the same level as the seated audience, and another four loudspeakers placed about 14 meters overhead (Figure 2).

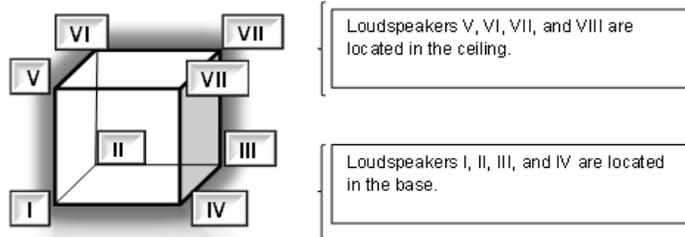


Figure 2, Consider that the console is placed in the centre of the auditorium, where the sound projectionist manually controls many of the spatial auditory events, and the audience seats around it.

The specific location of the loudspeakers, together with an array of different synthesizers, mixing consoles, and samplers, produces the peculiar effect of being in a 360° sound field, giving the impression of total motion. *Oktophonie* creates and develops four sound patterns or timbres resembling what we might term crashes, bombs, shots, and text-sounds, then combines them all together in some fashion. The sound patterns occur or “move” from vertex or take-off point simultaneously at different speeds in all sorts of directions: diagonally, from front to back, from side-to-side and from up-to-down, and rotationally in spirals to the right and to the left, with downward rotations from above to below—the alternating *tempi* gradually getting faster or slower, either *accelerando* or *ritardando*.

Each sound arrives at a different vertex or landing point among the loudspeakers or “moves” downwards or upwards corresponding to the initial sounds or glissando sounds—the sliding from one note to another. In a nutshell, sounds occur or “come from” all directions and rotate around the loudspeakers (I-VIII) with exceptional flexibility.

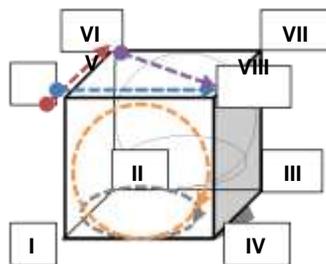


Figure 3, The sound patterns expand into the counterpoint in a three-dimensional soundscape. The cube’s organization allows sound “moving” from point to point--different arrows show how sounds “move” within this space.

The complexity of *Oktophonie*’s architecture seems to multiply on itself because of the effect of eight separate loudspeakers playing four different timbres combined and varied at different loudness. We are able to distinguish, however, these eight loudspeakers in each vertex (I-VIII) and within them aurally discern the four distinct timbres—the crash, bomb, shot, and text sounds—in at least some of their basic combinations. Moreover, we are able to aurally detect other properties and cues and to

estimate the distances among the sources simultaneously. Our auditory quality-space is multidimensional; it can map all these elements together. How does quality-space map mere sound locations<sub>p</sub> relative to *directions* in our experience of Oktophonie?

When loudspeakers V-VI-VII-VIII produce and encompass the different timbres at different speeds in diagonal directions, from front to back, horizontally and vertically, and in a spiral rotation from left to right (Figure 3), quality-space allows us to discern these distinct sound locations<sub>p</sub> and derivatively from that process, mapping their corresponding sound locations<sub>m</sub>. That is, one occurring from side-to-side (for example, maybe the crash-like sound produced by loudspeaker V and landing at loudspeaker VIII), another occurring from front to back (perhaps the bomb-like sound produced by loudspeaker V and landing at VI), another occurring in diagonal (the shot-like sound produced by loudspeaker VI and landing at loudspeaker VIII), etc.

The sound locations<sub>m</sub> are thus individuated by their position in a quality-space that we create relative to their sound locations<sub>p</sub>. Although an array of other audible properties are involved in an intimate relation with the identification of sound locations<sub>p</sub>, we have paid attention to mere sound locations<sub>p</sub>, so our auditory quality-space maps sound locations<sub>m</sub> from which we could describe our auditory experience.

### 3. Consciousness of Auditory Sensations and Conscious Thoughts

As mentioned (§1), when qualitative states are conscious one is conscious of being in those states by having thoughts—higher-order-states. As noted (§2), my view of an auditory quality-space captures the way in which we are able to discriminate among pitches<sub>m</sub>, timbres<sub>m</sub>, loudness<sub>m</sub>, and sound locations<sub>m</sub>. HOT theory, then, as it applies to qualitative mental states as against purely intentional states, depends on QST in order to describe our states in terms of corresponding mental qualities. That is, our higher-order-thoughts could capture qualitative distinctions of our having a particular auditory sensation exhibiting pitch<sub>m</sub>, timbre<sub>m</sub>, loudness<sub>m</sub>, and sound locations<sub>m</sub>.

HOT theory explains that a way in which we are conscious of our sensations may depend on attention and experience at discerning various mental qualities. The way of being aware of our states as being of different types can result from acquiring conceptual resources for the qualities of our sensations. In this section, I analyze how acquiring musical experience and conceptual resources for the qualities of our auditory sensations might affect the way we are conscious of those sensations with fine-grained mental qualities.

#### 3.1. Experience and Conceptual Resources

According to Rosenthal (2005: 186–7), ‘increased attention to one’s sensation seems by itself dramatically to enhance the way one is conscious of the qualities of the sensation’; furthermore, ‘repeated experience is also sometimes responsible for one’s sensations coming to seem to have more refined qualities’.

Let me elaborate: by paying attention more closely we can distinguish musical sound from incidental noise occurring at the same time. Depending on the amount of

repetition, constancy, and time spent in listening, our sensations can also seem to have more refined qualities: *Ceteris paribus*, the experience would not be the same for a hearer who attentively listens to Bruckner's Eighth Symphony twice a week for a year and for one who listens to it twice in a year. Both of them can achieve a certain awareness of their particular states with specific qualities, but they would be doing this in relatively differently refined ways. Every time one listens to a piece of music one could find different features of the whole work; with repeated experience, one's ear becomes more acute and sophisticated, and thus one could be aware of one's auditory sensations to a more complex degree (Renero, 2009).

Other repeated experiences and attention can also be relevant for our consciousness of the qualities of the sensation. We might be able to anticipate what would be played next, or recognize similarities and differences between parts of a piece or between different works (Davies, 1994: 326–31), or even distinct interpretations of the same piece.

One might question whether repeated experience and attention or much exposition could not impair fruition, and not only boost. To remain with the Bruckner symphony example: could not attention decrease dramatically after dozens of listening to it? Could not a decrement of attention imply a decrement of higher-order-thoughts?

First, my example is premised on a hypothetical listener who actually listens to music with interest. She is interested enough to listen to it repeatedly in order to learn something new every time that she hears to Bruckner's symphony. So every time she listens to this piece of music she can find different features—e.g., the listener can identify the accompaniment of *tremolando* strings in the scherzo, or the pulsating string chords in D flat over which violins have an expressive hymn-like theme in the adagio. Alternatively, at some subsequent time, she can notice the themes to which a movement returns, or the repetition of a phrase, as well as apprehend the theme's articulation, and so on.

Second, attention may be considered a necessary condition in order to gain access to aspects of her experience or to strength the way she has certain experiences. Also, since the listener is interested in the work, she pays attention on it. Certainly, once the listener loses interest in the work, her attention may decrease, and so the way she experiences the work in question. Whether a decrement of attention implies a decrement of higher-order-thoughts or not is a matter of a different discussion that I do not tackle here.

Now, let us see how acquiring conceptual resources for the qualities of auditory sensations can affect our consciousness of those states and the conscious thoughts we have about our musical experiences.

Although one could have higher-order-thoughts about qualitative character without language, within the HOT theory, acquiring conceptual resources for the qualities of one's sensations affects the way one is conscious of those sensations with *distinctive* qualities. This means that learning new words, their meanings and use, could result in new conscious qualities. 'One may not be conscious of any differences between the auditory sensations that result from hearing oboes and clarinets until one learns words for the different qualities. Learning those terms will then sometimes lead to one's being conscious of the sensory qualities as being of distinct types' (Rosenthal,

2005: 187). Again, Rosenthal does not elaborate on auditory experience; let us analyze it here.

Notice that possessing more language does not necessarily imply more capacity to discriminate: a demonstrative or observer may suffice, I think (as in mere sound location: “*that* sound..., *those* sounds...”). In order to have an experience with certain contents, the subject does not need to possess the corresponding conceptual capacity. Concept acquisition, however, may make an important difference in the subjective experience, and how things seem to us could ground the demonstrative reference in a distinctive (and specifically) recognized mental quality.

What is relevant is that, insofar as one acquires new language, presumably ‘a subjective difference emerges between mental qualities that had previously been subjectively indistinguishable’ (Rosenthal, 2004: 39; cf. 1997: 742).

But what is that difference? As I show (§3.2) the difference is in having *thoughts* with new contents; concretely, to have higher-order-*thoughts* that one is in an auditory state exhibiting certain auditory mental qualities. Thus, learning new words such as ‘oboe’ and ‘clarinet’ for the qualities of one’s auditory sensations can result in our having higher-order-thoughts that are more fine-grained with respect to distinctions among various mental qualities.

Let me suggest that the subject would learn the specific words for those qualities—mellow tones—and their sound sources—oboe and clarinet—, so she could also distinguish the source from the sound. So, concepts help us to individuate, to easily focus on something in particular, and to enhance our experiences.

If we again consider  $\text{pitch}_m$ ,  $\text{timbre}_m$ , and  $\text{loudness}_m$ , it would seem that learning a word such as ‘timbre’ for that character of sound we recognize as generated by an oboe versus a clarinet could make us aware of being in an auditory sensation with that distinctive quality. Once we acquire a refined vocabulary, we might become able to differentiate fine-grain mental qualities, and to discern consciously among them and our auditory states. By learning and using these words to label different mental qualities, then, we would be able to distinguish states with new qualities. Accordingly, our being able to have these qualities is not the result of a characteristic of the mental qualities themselves but rather of how our higher-order-thoughts describe that state in qualitative terms. I will go on now to show how this works in our musical experience.

### 3.2. Conscious Thoughts about One’s Musical Experience

Auditory quality-space posits mental qualities corresponding to discernible audible properties. However, with refined vocabulary and, especially, with experience, we could consciously discern among our auditory states that exhibit those qualities, and presumably come to have higher-order-thoughts that *describe* our auditory sensations in terms of the corresponding mental qualities. I conclude by demonstrating this view in action by commenting on a couple of musical fragments, and showing how a hypothetical listener could describe *and* report her thoughts in qualitative terms. We will see that describing things in auditory terms, plus the experience gained over time of discerning various auditory mental qualities, consists in having thoughts that organize the auditory experience.

Consider a person who has acquired the experience and concepts to discern timbres<sub>m</sub> and sound locations<sub>m</sub>. This person has the ability to point out when some of those qualities are present in her experience. In auditory sensing a fragment of Stockhausen's *Oktophonie*, the listener may describe back timbres<sub>m</sub> and sound locations<sub>m</sub> like this:

“Low rumbles, electrifying clashes, collisions, and crystal-like sounds explode simultaneously at the beginning. Then clearly distinguishing bomb-like and shot-like sounds appear from side-to-side. After that, bombs come from the lower loudspeakers whereas the shots come from the upper loudspeakers, and relatively slow changes to other timbres. Suddenly, these changes acquire a more noticeable character moving forward and onward with considerably energy. A bit later, a repeated and rapid sound succession paves its way through a horizontally moving spiral. It speeds up and slows down in both opposite ways”.

The listener here is describing the qualitative character of some timbres<sub>m</sub> in sound locations<sub>m</sub> comparatively in respect to directions or how different one sound location<sub>m</sub> is to another in quality-space.

Now, consider that this person has acquired the concepts of ‘pitch’, ‘dynamics’, and of their corresponding patterns of change. She discerns differences amongst dynamics<sub>m</sub> and the changes that go from loud to soft, and discerns pitches<sub>m</sub> from high to low within a passage, which she then applies in her descriptions of the qualities of her states. Although she has already aurally sensed and individuated the audible properties, by means of her higher-order-thoughts she can also be aware of being in those qualitative states with their distinctive dynamics<sub>m</sub> and pitches<sub>m</sub>. I will demonstrate how this might work by commenting on a fragment (*Allegro Moderato*) of Franz Schubert's 8<sup>th</sup> Symphony in B Minor, D.759. The listener describes her auditory sensations in terms of mental qualities:

0'04": Quiet and mysterious opening with low pitch (say, below other heard pitches), and very soft dynamics or *pianissimo*. 0'20": A middle range pitch. 0'28": A high pitch starts. 0'36": Sudden loud dynamic or *subito forte*. 0'52": The sound gets louder, a higher pitch is identified, and another change in dynamics: a long *crescendo* leading to *forte*, then *fortissimo*. 1'11": A sudden collapse of soft dynamics or *piano* followed by *diminuendo*. 1'21": A new tone, first low pitch, and then high pitch with marked *pianissimo*. 2'03": Finally, *fortissimo* again (Kerman, 2008: 9).<sup>18</sup>

Notice that the listener describes the qualitative character of the dynamics as well as the pitches themselves. In particular, she describes the dynamics<sub>m</sub> comparatively in respect to the relative graduation or how close one dynamic is to another dynamic<sub>m</sub> in quality-space. Yet it is possible that the listener is describing changes in the dynamics<sub>m</sub> as in 0'52", and pitches<sub>m</sub> within a range, and in respect to the antecedent in the temporal succession. The description of her auditory states consists in her having *thoughts* that organize her auditory experience of this passage.

Notice that auditory experience indeed exists in time: is temporally extended and has duration. When the listener hears this fragment, her auditory experience occurs in time—or at the most a few moments later. The listener is able to automatically incorporate changes, succession, and persistence. Conscious auditory states seem to occur in the same temporal framework as the auditory events happen, even though their

<sup>18</sup> Abbado (2010), Chamber Orchestra of Europe, Deutsche Grammophon.

exact timing is not easy to determine. In the previous example, I have marked time to represent the progression of the passage, and to show that in order for one's auditory sensations to be conscious, it is necessary they be conscious in respect of a relative time sequence with duration (this point will apply to the next example as well).

Now, the listener's experience of this piece would be enhanced if, during the hearing, she were to focus more on concepts and perhaps on technicalities. By 'experience' I mean the combination which results from our auditory states and those higher-order-thoughts whose contents are those states with fine-grained dynamics<sub>m</sub> and pitches<sub>m</sub>. By 'enhanced', I refer to the experience's being a richer variety of mental qualities, a way in which the listener's conscious thoughts might arise. In this sense, *ceteris paribus*, a hearer who identifies the various dynamics<sub>m</sub> and pitches<sub>m</sub> in Schubert's 8<sup>th</sup> Symphony would not be having the same experience as a hearer who hardly distinguishes between them.

Next, let me now return to Schubert's 8<sup>th</sup> Symphony and comment on approximately the same passage as heard by an experienced listener who has acquired a refined vocabulary for certain qualities, and can discriminate timbres<sub>m</sub>. She will be attentively listening to the passage and experiencing the music's flow with the timbre<sub>m</sub> as it changes over a period of two minutes, with the theme changing suddenly when it is played by different instruments. She will be having higher-order-thoughts that assert that she is in a state with these particular qualities. Since she has already acquired a refined language, these differences will be reflected in the intentional content of her thoughts, which are themselves consequently refined.

She might just express: "I hear x timbre<sub>m</sub>." However, I want to show that she would also be able to *report* the sensations she is having with distinctive timbres<sub>m</sub>. Notice that a listener's own reporting is an important way to be aware of her own current or recently past auditory experience.

Interestingly, the listener's reports provide evidence concerning her thoughts about how the musical experience appears to her: "At 0'04" *I think*<sup>19</sup> I hear some rustling-like double-basses sound opening the first movement. At 0'20" violins enter, and at 0'28" some woodwinds are incorporated into the strings. I think I hear clarinet-like mellow tones, while I barely hear a horn-like hollow sound. At 0'52" massed strings—like bright and shrill sounds—arise. I think that especially the violins sound powerful and intense; the sound of the cellos is gorgeous. At 1'11" a horns-like brassy sound is ringing, and a sudden collapse to smooth textures is being carried by strings again. I am not sure if there is a clarinet or a bassoon accompanying the strings. At 1'21" all the strings plays together; some of them seem to be trembling. I think I hear a cello-like plaintive sound and violin-like mellow sound. I hear undistinguishable woodwinds in the background. Finally, at 2'03" I hear a sharp cut off and then a sonic boom. I hear many instruments are playing together; they are hard to distinguish."

What is relevant here is that the listener is conscious of being in these different auditory states as they change sequentially. I have attempted to mark the time as a representation of reportability: the listener is able to report her thoughts or cluster of thoughts while she is conscious of the differences in her states with their respective mental qualities.

---

<sup>19</sup> Generally, though, we neither utter this mental attitude when reporting high-order-thoughts, nor are we so fine-grained in describing qualitative states.

Finally, HOT theory and QST in tandem can provide a potential experimental study to be developed. In the short term, I aim at giving an account of real time subjective auditory experience that could use reports and in turn illustrate some properties of our auditory experiences in connection to brain activity. Such an account could broaden the significance of the subjective experience by combining theoretical resources on consciousness and experimental results.

## Conclusion

In searching for an answer to the question of how our consciousness of auditory states arises, I have claimed that the rich and varied features of our auditory experience—specifically with regards to pitch, timbre, loudness, and spatial location—can be organized by our conscious thoughts, and all these make an important difference not only for our subjective auditory experiences *per se*, but also for what we process while thinking about those very experiences.

Using HOT and QS theories in tandem, and extending them to the auditory modality, I have shown that we can discern important differences among auditory mental qualities. I have further shown how experience and acquisition of conceptual resources can contribute to our being aware of the differences among our conscious auditory states. Finally, I have shown that our consciousness of auditory states consists in having thoughts that organize our experience of a musical passage.

HOT and QST gain further plausibility from their applicability to this domain, but my goal here is not simply to bolster these theories, but also to advance some specific proposals about auditory experience, which may prove useful to researchers with a range of theoretical orientations.<sup>20</sup>

## References

- Armstrong, D. M. (1968) *A materialist theory of the mind*, New York: Humanities Press.
- Blauert, J. (1997) *Spatial hearing: The psychophysics of human sound localization*, Cambridge, MA: MIT Press.
- Block, N. (2009) Comparing the major theories of consciousness, in Gazzaniga, M. (ed.) *The cognitive neurosciences*, Cambridge, MA: MIT Press.
- Block, N. (2011) The higher order approach to consciousness is defunct, *Analysis*, **71**(3), pp. 419–431.
- Brown, R. (2012) The myth of phenomenological overflow, *Consciousness and Cognition*, **21**(2), pp. 599–604.

---

<sup>20</sup> Earlier versions of this paper were presented in 2009 and 2010 at The Graduate Center, CUNY. Talks of this paper were given at The National Autonomous University of Mexico in 2011, Princeton University in 2012, and The University of Tokyo in 2013. I am especially grateful to Fermin Fulda, Rasmus Grønfeldt Winther, Yukihiro Nobuhara, John O’Dea, Rohit Parikh, Jesse Prinz, David Rosenthal, and Pedro Stepanenko for helpful comments. I am also grateful to the anonymous referees of JCS for motivating clarifications that improved this paper.

- Byrne, A. (1997) Some like it HOT: Consciousness and higher-order thoughts, *Philosophical Studies*, **86**(2), pp. 103–129.
- Carruthers, P. (2000) *Phenomenal consciousness: A naturalistic theory*. Cambridge: Cambridge University Press.
- Carruthers, P. (2005) *Consciousness: Essays from a higher-order perspective*, Oxford: OUP.
- Casati, R. and Dokic, J. (2009) Some varieties of spatial hearing, in O'Callaghan, C., & Nudds, M. (eds.), *Sounds and perception: New philosophical essays*, Oxford: OUP.
- Casati, R. and Dokic, J. (1994) *La philosophie du son*, Nimes: Jacqueline Chambon.
- Clark, A. (1993), *Sensory Qualities*, Oxford: Clarendon Press.
- Clark, A. (2000), *A Theory of Sentience*, Oxford: OUP.
- Davies, S. (1994) *Musical meaning and expression*, Ithaca: Cornell University Press.
- Dokic, J. (2007) Two ontologies of sound, *Monist*, **90**(3), pp. 391–402.
- Dretske, F. (1995) *Naturalizing the mind*, Cambridge, MA: MIT Press.
- Dretske, F. (1997) Conscious experience, in Block, N. et al., (eds.), *The nature of consciousness: Philosophical debates*, Cambridge, MA: MIT Press.
- Gennaro, R. J. (2006) Between pure self-representationalism and HOT theory of consciousness, in Kriegel, U., and Williford, K., (eds.), *Consciousness and self-reference*, Cambridge, MA: MIT Press.
- Gennaro, R.J. (2012). *The consciousness paradox: Consciousness, concepts, and higher-order thoughts*. Cambridge, MA: MIT Press.
- Hamilton, A. (2009). The sound of music. In O'Callaghan, C., and Nudds, M. (eds.), *Sounds and perception: New philosophical essays*, Oxford: OUP.
- Kerman, J., & Tomlinson, G. (2008) *Listen*, Boston: Bedford/St. Martin's.
- Kriegel, U. (2006) The same-order monitoring theory of consciousness, in Kriegel, U. and Williford, K. (eds.) *Consciousness and self-reference*, Cambridge, MA: MIT Press.
- Kriegel, U. (2009) *Subjective consciousness: a self-representational theory*, Oxford: OUP.
- Kulvicki, J. (2008) *The nature of noise*, Philosophers' Imprint, **8**(11), pp. 1–16.
- Lau, H, and Rosenthal, D. (2011) Empirical support for higher-order theories of conscious awareness, *Trends in Cognitive Sciences*, **15**(8), pp. 365–373.
- Lau, H., and Brown, R. (forthcoming). The emperor's new phenomenology? The empirical case for conscious experience without first-order representations, in Pautz, A. & Stoljar, D. (eds.), *Festschrift for Ned Block*, Cambridge, MA: MIT Press.
- Levine, J. (2001) *Purple haze: The puzzle of consciousness*, Oxford: OUP.
- Lotto, A., and Holt, L. (2010) Psychology of auditory perception, Wiley Interdisciplinary Reviews: *Cognitive Science*, n/a. doi: 10.1002/wcs.123
- Lycan, W. G. (2001) A simple argument for the higher-order representation theory of consciousness, *Analysis* **61**(1), pp. 3–4.
- Lycan, W. G. (2004) The superiority of HOP to HOT, in Gennaro, R. (ed.) *Higher-order theories of consciousness: an anthology*, Philadelphia: John Benjamins.
- Mandik, P. (2010) *Key terms in philosophy of mind*, New York: Continuum.
- Nagel, T. (1974) What is it like to be a bat? *Philosophical Review*, **83**(4), pp. 435–456.

- Nudds, M. (2001) *Experiencing the production of sounds*, *European Journal of Philosophy*, 9, pp. 210–229.
- O’Callaghan, C. (2007) *Sounds: A philosophical theory*, Oxford: OUP.
- O’Callaghan, C. (2008) Object perception: Vision and audition, *Philosophy Compass* 3(4), pp. 803–829.
- O’Callaghan, C., and Nudds, M. (eds.) (2009) *Sounds and perception: New philosophical essay*, Oxford: OUP.
- Pasnau, R. (1999) What is sound? *The Philosophical Quarterly*, 49(196), pp. 309–324.
- Pitt, I. (2010) Auditory Perception, [Online], <http://web.archive.org/web/20100410235208/http://www.cs.ucc.ie/~ianp/CS2511/HAP.html> [22April 2010].
- Prinz, J. J. (2012). *The conscious brain: how attention engenders experience*. Oxford: OUP.
- Randel, D. M. (1986) *The new Harvard dictionary of music*, Cambridge, MA: Belknap Press.
- Renero, A. (2009). Experience and consciousness: Enhancing the notion of musical understanding. *Crítica* 41(121), 23–46.
- Rosenthal, D. (1997) A theory of consciousness, in Block, N. et al., (eds.), *The nature of consciousness: Philosophical debates*, Cambridge, MA: MIT Press.
- Rosenthal, D. (2002) How many kinds of consciousness? *Consciousness and Cognition* 11(4), pp. 653–665.
- Rosenthal, D. (2004) Varieties of higher-order theory, in Gennaro, R. (ed.), *Higher-order theories of consciousness: An anthology*, Philadelphia: John Benjamins.
- Rosenthal, D. (2005). *Consciousness and mind*, Oxford: OUP.
- Rosenthal, D. (2010) How to think about mental qualities, *Philosophical Issues*, 20(1), pp. 368–393.
- Rosenthal, D. (2011) Exaggerated reports: reply to Block, *Analysis*, 71(3), pp. 438–443.
- Schnupp, J., Nelken, I., King, A., (2012) *Auditory neuroscience: Making sense of sound*, Cambridge, MA: MIT Press.
- Scruton, R. (1997) *The Aesthetics of music*, Oxford: OUP.
- Van Gulick, R. (2004). Higher-order global states: an alternative higher-order model of consciousness, in R. Gennaro (ed.), *Higher-order theories of consciousness*, Philadelphia: John Benjamins.
- Weisberg, J. (2011) Abusing the notion of what it’s likeness: A response to Block, *Analysis*, 71(3), pp. 438–442.