Musicology has a young disciplinary history, with regard to both its contents and its methods. After a first rough subdivision by Adler in a historic and systematic branch [1], there have been multiple attempts to broaden its scope with major new fields such as cognitive musicology, computational musicology, empirical musicology and others. Much of this early work, however, takes a rather disembodied stance toward music [2]. Hence, a whole new field of research has emerged in recent decades that proposes embodied and active models for the study of the musical mind as an ongoing process of dynamic interactivity between an organism and its environment [3–5].

The aim of this contribution is to put together different ways of knowing and processing music in a framework that allows us to think of them in a connected way. It argues for a back-to-basics approach to ways of apprehending that goes beyond opposed points of view and conceives of music as an epistemic category. Music, in this view, is to be considered as a sound-time phenomenon that has the potential of being structured by music listeners, with levels of processing that range from direct reactivity to more elaborate reactions to the sounds. What matters, then, is a way to “apprehend” music, rather than merely “perceive” it, with listeners being aware of a number of complementary ways of sense-making that constitute an assemblage of responses, deployed in relation with each other.
STARTING DICHTOMIES

Taking an epistemological stance toward music entails a definition of music as a temporal and sounding art. This means that music cannot be described in a disembodied way as “something out there,” but rather as something that impinges upon our senses in an actual “here and now.” Musical sense-making, on the contrary, is a kind of ongoing knowledge construction with a dynamic tension between actual or real-time sensation and mental representation of sounding events that are the result of previous experiences with meaningful sounds—a kind of acquired knowledge base—and the assignment of semantic weight to those sounding events that are the subject of ongoing attention. Real-time sensations, in this view, are transformed in representations that build up during the listening experience proper, a process that relies on short-term memory with a transition to semipermanent representations to be stored in long-term memory. The question can be raised, however, as to the distinction between memory and representation. Memory is factually driven in the sense that it implies a kind of reenactment of something that has been heard before. Representations, on the contrary, have a certain distance with respect to the actual sounding stimuli. They exhibit a more virtual character as mental replicas of sounds, which makes it possible to combine them in mental imagery in ways that have not yet been heard before. The use of the term representation, however, can be disputed to some extent, as evidenced by the recent shift from the traditional representation-centered framework toward a paradigm that views cognition as active, skillful activity that involves ongoing interactions with the world [8–11]. As such, there are four major dichotomies that may be considered as basic starting points for making sense of music: (i) the focal versus synoptic overview, (ii) the continuous/discrete distinction, (iii) the distinction between sensory experience versus cognitive economy, and (iv) the in-time/outside-of-time distinction.

The focal versus synoptic overview is related to the dynamics of temporal representation [12], which can be considered as the scanning of the sonorous unfolding in a global or synoptic way as against the sequential succession of focal acts of attention. Crucial in this distinction is the temporal window through which the listener keeps track with the unfolding of time [13]. This is constrained in the act of perception and has been coined the psychical present [14] or specious present [15] by Stern and James, as a moment in time that sharply divides past from future but is clearly distinct from both. It is possible, however, to extend this temporal window by recollecting previous sounding events in memory and by anticipating future sounds in imagery. This is the synthesizing function of imagery, which provides a simultaneous representation of multiple sounds, making it possible to conceive of larger spans of time at a glance [16].

The discrete/continuous dichotomy finds its foundation in signal processing theory. Discrete, in this context, refers to things or events with unit character that are clearly separated from each other. Such a discrete approach reduces the continuous flow of sounding music to the allocation of meaning to focal points in the sonorous articulation and conceives of music as a succession of encapsulated, discrete things that seem to exist also outside the time of actual unfolding. It calls forth conceptualizations that can be represented at an abstract-symbolic level in imagery, allowing listeners to manipulate mental replicas of the sounds in a kind of virtual acoustic space. The continuous approach, on the contrary, is perceptually bound and proceeds in real time with the listener keeping step with the music in a continuous and ongoing process of epistemic interactions with the sounds [17]. As such, the distinction is related to the digital/analog distinction: The digital approach conceives of separate and distinct units that can be addressed in an all-or-none way (e.g. yes or no, 0 or 1; hence the term digital); the continuous approach is analogous to the real world (hence the term analog) and portrays the minute particularities of the temporal unfolding. The distinction is exemplified in Fig. 2, which depicts a score notation and waveform representation of the first eight bars of the second movement of Schubert’s String Quartet no. 14. The upper pane shows a discrete-digital representation, using note symbols, which can be recognized and conceptualized as distinct elements with unit character; the lower pane shows an analog-continuous representation, with an attempt to discretize one of the two stereo tracks by drawing circles on top of it. It is up to the listener to delimit the scope of the discrete units—in this case either a half note or two quarter notes—though such parsing occurs mostly automatically. It is possible, however, to discretize at will and to rely on deliberate and focal attention that is driven also by personal choices and not merely by the structure of the music. The symbolic notation is easier to follow, whereas the continuous representation provides more information with regard to the temporal and dynamic unfolding.

The digital/analog distinction is related to the distinction between sensory experience and cognitive economy. Analog-continuous processing is time-consuming and proceeds in real time; conceptual knowledge is needed to manage information in a more economical way. The analog-continuous approach, then, is more sensitive, as it is not constrained by the limitations of fixed thresholds and distinctions, which are so typical for the discrete-digital approach. It is also closer to the real world, which presents itself not in a segmented way but in continuous transitions. The continuous way of decoding, therefore, is most suitable for “exploring” and “perceiving,” as advocated in the experiential approach to cognition [18]; the conceptual way, on the other hand, is more reductive in the sense that it constrains the real world from a relatively large and continuous set of stimuli to a finite set of discrete elements. This is a process of digitization that takes a piece of information from a richer matrix of information in the sensory-analog representation to feature it to the exclusion of everything else. Yet, in passing from sensory experience to conceptual representation, there is a kind of stripping away of components of information, which reduces the experience of the phenomenally rich thing to only one or some of its components [19]. This allows listeners to process information
with less cognitive effort, but at the cost of neglecting the idiosyncrasies of each particular rendering in order to allow discrimination at a more abstract level of similarity and to “recognize” things rather than to “experience” them.

The in-time/outside-of-time dichotomy, finally, has as a critical distinctive feature the presentation to the senses [20] (as I have discussed previously [21]). In-time processing involves the sensorial aspect of capturing sound with its impingement on our senses. It is characterized by consumption of time and perceptual binding, with a kind of continuous scanning of the temporal flow in an ongoing succession of acts of focal attention. As such it taps the moment-to-moment history of epistemic interactions with the sounds. As soon, however, as there is a kind of discretization of the temporal unfolding, by assigning semantic weight to focal points or limited temporal windows, there is also a beginning of outside-of-time processing, as the labeling and conceptualization refer to concepts that exist also outside of the time of actual unfolding. This allows listeners to go beyond the narrow temporal window of the actual now-moment and to mentally navigate through the music in a kind of virtual musical space. This is, in fact, the computational approach, which places itself outside the time of actual unfolding.

TOWARD A FUTURE EPISTEMOLOGY: THE COMPUTATIONAL AND THE DYNAMIC APPROACH TO MUSICAL SENSE-MAKING

The above-mentioned dichotomies reflect to some extent the distinction between a “computational” and a “dynamic” approach to musical sense-making and cognition in general [22].

The computational approach relies mainly on discrete sequences of static internal states with symbolic representations of stored or represented knowledge without relation to the actual sounding in real time. It makes it possible to perform mental arithmetic on these representations and to process concrete-sounding sonorous events not as sensory data but as data that are disengaged and detached from any particular here-and-now situation. Such an approach deals with symbols that function as mental replicas of the sounds, allowing a kind of symbolic play, which is characterized typically by the plasticity and reversibility of mental operations. By keeping distance from the perceptual flux, it reflects the influence of human linguistic capacity on music cognition, which allows people to understand experiences via labels or names, not only via direct involvement. Music cognition, then, is considered in terms of recognizability, prototypicality and communicability, stressing a shared experience rather than the perceptual qualities. It is an approach that is typical of the logocentric or logogenic approach to musical sense-making [23], which states that musical meaning is conducive to verbal description or symbolic labels.

The dynamic approach, on the contrary, emphasizes the role of ongoing (epistemic) interactions between an agent and its environment. It stresses the fundamental role of the immediate situation and the primacy of acting in the actual world with agents capable of carrying out open-ended tasks in unconstrained environments. As such, it unfolds in real time with the nervous system, the body and the environment continuously evolving and influencing one another [24]. The theoretical framework for this approach refers to the language of dynamical systems theory, which revalues somewhat
some of the operational terminology of behaviorism. What matters in this approach are the major moments of perceiving, acting and perception-action coordination as well as the feedback loops that are needed for effective control of acting on the environment. The “circularity” of stimulus and reaction, therefore, is central in epistemic interactions with the world at large and with the sounding music in particular. Dynamic approaches to cognition, however, are not exhausted by cycles of stimuli and responses but postulate a much more complex unfolding dynamics of reciprocal interactivities [25].

The translation of these insights to the realm of music, however, is still in its infancy. It calls forth the transition from mere sonic experience to musical sense-making with a dynamic tension between the experiential-perceptual and symbolic-conceptual approach to music cognition [26]. Consider, for example, the art of castrates in the eighteenth century, with male sopranos showing a level of virtuosity and singing technique. Comparing their actual acoustic rendering with the score notation shows immediately the simplistic and highly reductive character of the note symbols, which are merely hints and triggers for the singer. Figure 3 gives an iconic example, showing the score notation and spectrogram of the first two words from the aria “Alto Giove” from Nicola Porpora’s opera Poliferno, as documented in the film Farielli, Il castro. It is intriguing also to compare performances by distinguished countertenors such as, for example, Derek Lee Ragin, Philippe Jaroussky and Franco Fagioli, to see the idiosyncratic particularities of each of them.

It can be questioned, further, whether we could conceive of a dynamic definition of musical sense-making that does justice to a conception of musical events as higher-order variables that can be defined as continuous functions of time that have the possibility of discrete denotation [27,28]. Dealing with music, in this view, is sensory-driven but amenable to symbolic computations with a gradual shift from presentational immediacy, with information being presented to the senses, to symbolic representation in the absence of sensory stimulation. The question can be raised, however, whether it is at all possible to conceive of real symbolic representation in the case of music, with its predominantly auditory component. There are, in fact, immediate and lifelong interactions with musical stimuli, which may result in lasting changes in the internal state of the listener [29]. Such changes are not static symbolic representations but are dynamic and ongoing. They are the result of knowledge construction as the outcome of previous and actual interactions with the sounds. What I argue for, therefore, is a dynamic and constructivist approach to musical knowledge and information, not as disembodied but as embodied and situated material that is dealt with in a dynamical way [30]. These claims have not yet been stated explicitly in the literature, though there were some phenomenological approaches to the analysis of music in the ’80s that revolved around the conception of “music as heard” and “music as experienced,” i.e. as something that has existential structure and meaning [31,32]. It means that our actual involvement with music is experienced rather than being solely reasoned and interpreted [33], or put in other terms: It is drastic rather than gnostic, to use Jankélévitch’s terms [34,35]. The consequences of this approach are considerable, calling into question the status of the score, the role of our theoretical notions about music and the analytical methods to be used to study music as a sounding phenomenon.

Fig. 3. Score notation and spectrogram of the first two words from “Alto Giove” from Porpora’s opera Poliferno. (© Mark Reybrouck)
References

12 Rolf Inge Godøy, Formalization and Epistemology (Oslo: Scandinavian Univ. Press, 1997).
27 Reybrouck [17].

Manuscript received 9 July 2019.

MARK REYBROUCK is emeritus professor at the University of Leuven and guest professor at Ghent University. His research agenda concerns musical sense-making with a major focus on the biological roots of musical epistemology and the embodied and enactive approach to dealing with music.