Chapter 2

From Sonic Environment to Soundscape

Dick Botteldooren,[1] Tjeerd Andringa,[2] Itziar Aspuru,[3] A. Lex Brown,[4] Danièle Dubois,[5] Catherine Guastavino,[6] Jian Kang,[7] Catherine Lavandier,[8] Mats Nilsson,[9] Anna Preis,[10] and Brigitte Schulte-Fortkamp[11]

[1]Ghent University, Ghent, Belgium

[2]University of Groningen, Groningen, The Netherlands

[3]TECNALIA Research and Innovation, Bilbao, Spain

[4]Griffith University, Brisbane, Australia

[5]Pierre and Marie Curie University, Paris, France

[6]McGill University, Montréal, Canada

[7]University of Sheffield, Sheffield, United Kingdom

[8]Université de Cergy-Pontoise, France

[9]Stockholm University, Stockholm, Sweeden

[10]Adam Mickiewicz Univeristy, Poznań, Poland

[11]Technical University of Berlin, Berlin, Germany

CONTENTS

2.1 Listening 19
- 2.1.1 Attentive, Analytic, Descriptive Listening: Most Popular in Soundscape Research 19
- 2.1.2 Holistic Listening and Hearing: A More Hidden Contribution to the Soundscape Experience 20
- 2.1.3 Different Listening Styles: All Part of the Same Experience 20

2.2 Auditory Scene Analysis 20
- 2.2.1 Auditory Scene Analysis and Familiarity with Sounds 21
- 2.2.2 Effects of the Complexity of Auditory Scene Analysis on Soundscape 22

2.3 Role of Attention 22
- 2.3.1 Saliency-Driven Attention 23
- 2.3.2 Multisensory Attention 23
- 2.3.3 Attention to Location 24
- 2.3.4 Voluntary (Endogeneous) Attention and Attention to Memory 25

2.4 How Sounds Get Meaning 25
2.4.1 Meaning and Associative Memory 26
2.4.2 Linguistic Discourse as an Expression of Meaning 26
2.5 Appraisal and Quality Judgement in a Sociocultural Context 27
2.5.1 Appraisal and Affective Neuroscience 27
2.5.2 Appraisal and Predictability 29
2.5.3 Appraisal, Coping, and Behavioural Options 29
2.6 Holistic Model for Soundscape 30
2.6.1 From a Use Case Perspective 30
2.6.2 From the Perspective of a Sonic Environment 32
2.7 How Soundscape Theory Can Affect Practice 33
2.7.1 Measuring Soundscape 33
2.7.1.1 Measuring People 33
2.7.1.2 Measuring with People 34
2.7.1.3 Measuring with Computers in a Human-Mimicking Way 35
2.7.2 Soundscape Design 35
2.7.2.1 Designer's Vision and Possible Use of the Place 35
2.7.2.2 Composition 36
2.8 Future Directions and a Way Forward 37
References 38

Chapter 2 focuses on creating deeper understanding of the relationship between the sonic or acoustic environment and soundscape. It attempts to relate the somewhat vague concept *soundscape* to findings from psychophysics, psychology, hearing system physiology, and auditory cognition. The term *soundscape* has been used by different communities of practice (e.g., acousticians, composers, architects, ecologists, and psychologists), giving rise to several definitions (see Chapter 1). A standardized definition may not be required, but it is useful to summarize generally accepted views on this concept:

- The soundscape is evoked by the physical sound environment—henceforth called the *sonic environment*—but it is not equal to it, and therefore cannot be measured using classical sound measurement equipment alone.
- The soundscape is formed within a context. This context is shaped by all sensory stimulations—of which visual observations are most important—and by the knowledge people have accumulated about the place, its use, its purpose, its cultural meaning, their own and others' motivations and purposes to be there, the associated activities, and so forth.
- The soundscape concept tends to be used mostly in relation to open outdoor places, but it also has applications for indoor settings, mainly

public, but also private. It always entails a sense of spaciousness. Environmental sounds intruding in private space result in effects following different mechanisms, with control as an important factor.

- The timescale related to soundscapes is in the order of minutes to hours. The quality of the soundscape in some parts of the living environment can nevertheless have long-term effects on the quality of life and health of the population (see Chapters 3 and 5).

The following sections discuss various aspects of the relationship between the sonic, or acoustic, environment of a place and the person experiencing the soundscape to finally construct a holistic model. After a brief discussion of listening styles, this chapter will focus on low-level auditory scene analysis that leads to the formation of auditory streams and objects. It will then continue by introducing the important role of auditory attention in selecting which of these auditory objects will be noticed. Attended sounds get meaning, so this important aspect of soundscape analysis and design will be the focus of the next section. Finally, appraisal and quality judgement will be addressed as the last step in the process. Once our current knowledge on each of these aspects has been thoroughly investigated, a more holistic view on the relationship between the sonic environment and soundscape will be discussed. The last section of this chapter will highlight how this knowledge can be applied in practice for measuring and designing soundscape. This last part could be seen as an introduction to the next chapters in this book.

2.1 LISTENING

Listening is a complex process that involves multileveled attention and higher cognitive functions, including memory, template matching, foregrounding (attentive listening), and backgrounding (holistic listening) (Truax, 2001). Some scholars group listening styles in everyday listening and musical listening, thereby focusing on an apparent difference between music, a sound that is produced for a purpose, and all other sounds (Gaver, 1993a, 1993b). Everyday listening in this terminology focuses on the sound source, musical listening of the sound itself.

2.1.1 Attentive, Analytic, Descriptive Listening: Most Popular in Soundscape Research

In investigations where persons are asked about their aural experience in a place, researchers found that these persons most often mention particular sounds by naming the source of these sounds (McAdams, 1993). One could conclude from this that attentive, analytic, descriptive listening is the most important listening style in relation to the soundscape experience.

This is, however, only partly true. Asking visitors of a place to describe their listening experience automatically triggers attentive and descriptive listening. In absence of the researcher, this listening style would only be important in those special cases where the intended activity includes a strong attention focus on the environment or when the sound is so prominent and salient that listening to it cannot be avoided. Even musical listening, although it focuses on the sound rather than on its sources, can still be regarded as a type of attentive listening.

2.1.2 Holistic Listening and Hearing: A More Hidden Contribution to the Soundscape Experience

One should not underestimate the potential role of holistic listening or even simply hearing as a mediator in creating mood and appraisal of the sonic environment. Sound not actively attended to, and thus pushed to a background, can still have meaning (see further).

To our knowledge, there is no research that directly relates soundscape listening to the cognitive effort it requires. However, as attention assigns more cognitive resources to a sensory input stream, it is reasonable to assume that attentive listening, and in particular analytic, descriptive listening, requires more cognitive effort and is, as a consequence, also slower. Holistic listening and backgrounding frees cognitive resources for other tasks that might be more relevant at this instant in time. Yet, holistic listening is also expected to be faster, and thus allows the organism to more quickly create a mental image of its environment and act correspondingly.

2.1.3 Different Listening Styles: All Part of the Same Experience

As the listening experience in a sonic environment evolves, the listener switches between different listening styles: from the more holistic listening in readiness, waiting for familiar or important sounds to emerge (expected or not), to listening in search, expecting particular sounds in a context, or even to narrative or story listening—musical listening could be seen as a specific example of this listening style—focusing attention on one particular sonic story within the multitude of sounds.

2.2 AUDITORY SCENE ANALYSIS

The sonic environment of interest in the context of soundscape consists of a multitude of individual sounds. One of the first tasks of the auditory system is to analyze this auditory scene and identify its building blocks, a process referred to as auditory scene analysis (ASA). ASA involves decomposing a complex mixture of incoming sounds, originating from different sources, into individual auditory streams, using different auditory, but also visual and other, cues (Bregman, 1994). Auditory streams are classically regarded as existing in a preattentive phase. Although this view is appealing because of its conceptual simplicity, recent findings suggest that attention also plays a role in the formation of auditory streams (Cusack et al., 2004; Shamma et al., 2011). Overall, it can be stated that the process of auditory scene analysis draws on low-level principles for segmentation and grouping, but is fine-tuned by selective attention (Fritz et al., 2007). Sound objects within the sonic environment are thus formed with the help of selective attention (attention mechanisms will be explored further in the following paragraphs). In relation to the discussion in the previous paragraph, this implies that even in holistic listening, stream and object formation occur, yet they may be less precise than during attentive, descriptive listening.

Scene analysis is partly multisensory, although the relative importance of the components of this scene analysis in vision and audition is different. In auditory scene analysis, temporal grouping at timescales from seconds to minutes or even hours is extremely important. Grouping at shorter timescales is more likely to occur—at least as a first estimate—during the preattentive phase. Spatial cues obtained through binaural hearing also help the stream segregation process. Binaural unmasking is known to be a key factor in targeted story listening within a masking background noise. With environmental sound and soundscape in mind and with the relative importance of different listening styles discussed above, binaural cues may be less significant for stream segregation in this listening context.

2.2.1 Auditory Scene Analysis and Familiarity with Sounds

Identification of auditory objects based on spectrotemporal features is a learned process where learning relies on co-occurence of these features. The importance of temporal coherence in auditory scene analysis and learning in humans has recently been confirmed on a neurological basis (Shamma et al., 2011). As such, the familiarity of the listener with a sound may contribute to the ability to distinguish this sound in a complex sonic environment. Prior experience could therefore even have an influence on this low-level preattentive ASA and lead to interindividual differences in perceiving the sonic environment and in the soundscape experience. This ASA skill is transferable between sounds, as the ability to group and identify features may be influenced by early sound experience of any style: language, music, or even the simplest sounds in the daily living environment. This could also lead to cultural differences, as typical language and musical sounds may differ across cultures.

2.2.2 Effects of the Complexity of Auditory Scene Analysis on Soundscape

In general, one could expect that a sonic environment where various auditory streams can easily be formed is appreciated as a high-fidelity soundscape. More complex situations that cannot easily be "read" by the average listener may be perceived as too complex and mentally stressing. Like in the attention restoration theory (ART) and the associated fascination, urban environments, with too complex stimulations, could be a source of attentional fatigue (Payne, 2013; Kaplan and Kaplan, 1989). If, on the contrary, ASA results in a single auditory stream, the sonic environment may be perceived as boring.

The ability to segregate sounds from complex mixtures differs between persons, as does the amount of cognitive resources needed for this task. Aging, for example, does not affect the ability for sequential streaming, but it has a pronounced effect on concurrent sound segregation (Snyder and Alain, 2007). These interindividual and age-related differences in even this lowest level of processing may lead to differences in the degree of complexity in the sonic environment that is desirable.

2.3 ROLE OF ATTENTION

Let us now focus on the attention mechanism in more detail. The role of selective attention is to allow part of the sensory input to be evaluated in the context of specific knowledge while preventing sensory signals from overloading the higher-level cognitive system. Overall perceptual load thus plays an important role in attention mechanisms (Lavie et al., 1994; Lavie, 2010). In situations where soundscape analysis and design are usually applied, the use of the place does not focus on communication between a performer and a group of listeners (such as a musical performance or theatre). Apart from the verbal communication they may be involved in, most users of the public space have little interest in listening for particular sounds, such as birds or insects. However, as the auditory system always stays alert, sounds within the sonic environment could draw attention. The proposed theoretical model foresees a two-stage mechanism to account for this: auditory stimuli may draw attention because of specific features they possess, but they don't necessarily get attended. This two-stage mechanism is supported by neuroscience: sounds with high saliency trigger early brain response (Escera et al., 1998), while inhibition of return (Prime et al., 2003) and voluntary attentiveness to sound determine whether a late response corresponding to actual attending is observed.

2.3.1 Saliency-Driven Attention

Identifying sound features that increase saliency (Kayser et al., 2005) and attract attention is an important aspect of the proposed soundscape theory. It is well known (Kayser et al., 2005) that spectral and temporal variations and modulations—sometimes referred to as ripple—increase saliency for human observers. However, the auditory brainstem, which is responsible for these specific sensitivities, has a much higher plasticity than originally thought. On the basis of this, one could expect a common basis for auditory saliency, but in addition, some specificity for different (groups of) people.

Saliency sound features have been implicitly used in earlier studies in other fields, such as noise annoyance and sound quality. Tonality (Hellman, 1982), rhythm, or periodicity and impulsiveness have been introduced to explain differences in annoyance caused by different (industrial) environmental sounds (ISO 1996-1, 2003). Likewise, sharpness, roughness, and relative approach (Genuit and Fiebig, 2006) describe features of the sound that attract attention. And, of course, the loudness of the sound itself is an important factor in its saliency. Event-related loudness—where all other saliency features are kept constant—even in a complex and distracting environment, explains most of annoyance (Sandrock et al., 2008).

The saliency mechanism could also be evoked to explain observations at a more abstract level of auditory processing. Sounds—or their nomic or symbolic mapping—can also be called salient if incongruent with the context. Such saliency could draw attention to the symbol rather than to the sound features. In the latter case, incongruence of the sound in the scene can enhance detectability (Gygi and Shafiro, 2011). Event-related potential measurements confirm the deviant processing, also with complex sounds, but also show that familiarity with the sound has an effect (Kirmse et al., 2009). A foundation for rapid extraction of meaning from a familiar environmental sound was observed even when sounds were not consciously attended. Outward-oriented mechanisms in turn draw attention to the sound features corresponding to the symbol (see further).

2.3.2 Multisensory Attention

The listener embedded in a real environment—in contrast to experimental conditions—relies on all senses to structure a representation of the environment (Driver and Spence, 1998). One sensory modality could also draw spatial attention to a different modality and even strongly influence the perception itself. This raises the question of whether attention resources are controlled by a supramodal system or by many modality-specific attention systems. In focused attention conditions, judging each signal (sound and vision) separately when incongruent sensory signals occur at the same location is difficult, at least much more difficult than when the incongruent signals come from different spatial locations and attention is divided (Santangelo et al., 2010). A multilevel mechanism of attention with a multimodal component overarching the single sensory component seems the most plausible model given today's knowledge. In the context of assessing the sonic environment, this could be interpreted as a stronger emphasis on visible sources, but at the same time a lower identification probability of a deviant sound experience if this sound comes from the same location as the visual stimulus.

Multisensory attention mechanisms also have a strong temporal component. Sound stimuli presented in temporal congruence with the appearance of a visual target make the visual target pop out of the scene (Talsma et al., 2010). Likewise, visual stimuli that appear—independently of where they appear—at the same moment that a sound could be detected increase the probability that attention will be paid to that sound.

Based on this knowledge on multisensory perception, a long-standing concern of soundscape designers can at least be partly answered: Is it advantageous to hide unwanted sound sources from view? From the attention perspective, one could conclude that provided that the sound is not very salient, and thus is not very likely to attract attention, noticing the sound can be avoided by eliminating visual stimuli that are congruent is space and time with the unwanted sound. Similarly, a wanted sound should be accompanied by a visual stimulus to ensure that it receives proper attention. It should, however, be noted that in case of very salient sounds that will certainly attract attention, the absence of a visual stimulus may come more as a surprise, which may influence appraisal.

2.3.3 Attention to Location

This brings us to the point of binaural hearing. Inhibition of return on location (Mondor et al., 1998) could explain why moving sources or groups of sources of the same kind popping up at different locations might be less easily inhibited by the auditory system and thus continue to attract attention longer than a stationary source. It is known that identity information predominates over location information in auditory memory (Mayr et al., 2011); thus, soundscape appraisal (see the following sections) in itself—in contrast to unmasking—may be less sensitive to aspects of binaural hearing.

Source–listener distance is another aspect of location that might influence attention. As loudness is a primary clue for distance perception, and loudness—or at least loudness change—is known to influence saliency and thus attract attention, there seems to be indirect evidence that sounds from close-by sources would attract more attention. However, we found no experimental evidence that nonfluctuating sounds from a source at a close distance would attract attention more strongly than louder sounds at longer distances.

2.3.4 Voluntary (Endogeneous) Attention and Attention to Memory

Listening in search or story listening involves voluntary (endogeneous) attention focusing grounded in higher-level cognition. It can be shaped by expectations about the place based on prior experience or knowledge, or it can be initially triggered by involuntary attention focusing. Familiarity with the sound is a prerequisite for voluntary attention focusing, yet unfamiliar or incongruent sounds are more likely to attract attention for reasons that can be explained as complex saliency (see above). The interplay between involuntary and voluntary attention results in sustained attention to particular sounds in the sonic environment. Known sounds—which are most likely sounds with strong meaning—could therefore easily be used in soundscape design to draw and maintain attention. Using unfamiliar sounds—as an element of surprise—may need additional context, visual, for example, to ensure that they get attention.

Occasionally, intended activities—and the reason to go to a place—involve listening in search or story listening. For example, one could expect that voluntary attention is focused on natural sounds (birds, breaking waves) if a person is visiting a place to experience nature. However, the complex interplay between expectation, appraisal, and coping may also lead to increased attention focusing on the unexpected or unwanted sound.

The reaction of the brain to sensory stimuli depends on its current state. According to the attention to memory model hypothesis, very similar attention mechanisms are involved in memory tasks, on the one hand, and sensory processing tasks, on the other (Cabeza et al., 2011). Part of the neural circuitry even seems to overlap. This implies additional modulation of overall attention devoted to the sonic environment. Conversely, it also implies that sensory input in general and sound in particular can distract from memory (and cognitive) tasks. Soundscape perception can therefore be different for the same person at different instances, dependent on current activity.

2.4 HOW SOUNDS GET MEANING

The role of audition is not mere information processing but recognition, resulting from bottom-up (signal-driven) and top-down (knowledge-driven) processing. The knowledge-driven component should not be underestimated. The sensory perception could even be regarded as a factor correcting and fine-tuning the mental representation of the (sonic) environment. As such, the meaning of sound(s) could be determined as much by the current state of the mind (emotions included) as it is by the stimulus per se.

2.4.1 Meaning and Associative Memory

Meaning to a large extent depends on the associations a stimulus evokes. The process of attaching meaning to (components of) the sonic environment includes several stages of abstraction. At the lowest level, the association between the sound objects and events they stand for is activated. In case of nomic mapping, the sound and events present consistent information. The event itself produces the sound, for example, a car approaching. In this case, the sound source (e.g., the car) will often be the most important factor in creating meaning. Symbolic mapping relates a sound to an event that does not produce the sound. Symbolic representations allow sharing individual experiential meaning and contribute to the elaboration of social (conventional) meaning. For example, church bells are mapped to an event that is not the ringing of a bell per se, but a socially defined event, such as celebration, in a specific culture.

Meaning extends to more abstract levels by associating the recognized event or source to a larger set of concepts in a somewhat vague way. This vague meaning is sharpened by knowledge of the place (Niessen et al., 2008) and by the most recent meaning attached to the auditory stream. The latter could explain why the path followed by the person experiencing a sonic environment may influence the interpretation and appraisal of a sonic environment.

In order to understand the meaning given to the (components of the) sonic environment, one thus has to understand how associations are learned. An organism learns in order to better predict, prepare for, and anticipate possible futures based on the current situation (Bubic et al., 2010) and to evaluate behavioural options. For the current discourse, a few elements from the multitude of learning theories are extracted. Learning can occur because of prediction error or a teaching signal. From a neurological point of view, there might not be that much difference between both types of learning, but from a sociological point of view, they have very different consequences. For learning from prediction error, the organism must have "lived" the consequence of a wrong prediction of the events occurring in its environment. This makes this type of learning different for different persons, although many of the events that people experience are very similar within a given culture, geographical area, and given generation. However, the influence of culture, geographic area, and generation becomes even more pronounced while learning from peers. Thus, although associative memory is individual, some common features can be expected.

2.4.2 Linguistic Discourse as an Expression of Meaning

Although some scholars may argue that meaning and verbal description are very closely related and can thus be unified, a small distinction remains that could clarify some observations made in soundscape research.

The linguistic label assigned to a sound or the event or source it stands for relies on a complex process, such as categorization and naming, where one category does not depend exclusively on its intrinsic properties, but also on its resemblance to and differences from other categories within the whole classificatory system. Therefore, the same signal can be categorized at different levels of specificity (e.g., traffic, car, sports car; human voices, child voice, my child calling) or along different principles of categorization (source, event such as car breaking or starting, global appraisal, wanted vs. unwanted sound). Thus, not only the context of the observation seems to matter, but also the context in which the meaning of a sound is expressed linguistically.

To complicate matters, the meaning of a word or linguistic expression can also differ between persons. A typical example of relevance in soundscape research is description of tranquility and tranquil area. The French word *calme*, for example, was found to represent different things, depending on the persons asked: a more social interpretation with matching human vocalizations, an evocation of nature and natural sounds, or a notion of quietness and absence of sound (Delaitre, 2013). Thus, a match has to be found between the meaning of words and expressions, on the one hand, and the meaning the sonic environment evokes, on the other hand, in order to understand how persons describe their soundscape experience using a narrative.

2.5 APPRAISAL AND QUALITY JUDGEMENT IN A SOCIOCULTURAL CONTEXT

Appraisal and quality judgement of soundscape form the final step in the analysis. These processes can be regarded from different perspectives. Three different perspectives are discussed below: the perspective of neuroscience reveals how low-level brain functionality could explain why certain sounds are appraised more positively; the perspective of learning and predictability could explain the influence of expectations as well as the preferred level of complexity; and the perspective of coping and behavioural options allows us to view appraisal of the sonic environment in its context in the most holistic way.

2.5.1 Appraisal and Affective Neuroscience

Brain imaging techniques are increasingly used to study how classical psychological concepts are encoded in the brain and to identify connectivity and causal relationships between activation of different brain areas. For the discourse of appraisal and quality judgement, the reward system seems worth looking at. Neuroimaging studies have found that the affective valence of pleasure may be coded in separate networks of brain areas from sensation intensity. The reward system can be described as adding hedonic gloss to the sensation, which could be experienced as conscious pleasure (Berridge and Kringelbach, 2008).

Reward comes in different flavours (Berridge and Kringelbach, 2008): liking, wanting, and learning. Liking is the actual pleasure component of award. At the first level, core liking reactions occur that need not be conscious; at a second level, cognitive brain mechanisms of awareness may elaborate conscious liking from this. Wanting is a motivation for reward that could be a conscious desire to reach cognitive goals, but it also has an incentive saliency component that is not necessarily conscious. Wanting can be, but is not necessarily, linked to liking. Learning includes expectations about future rewards that are learned by association, representation, and prediction. The process can be the result of explicit cognitive reasoning, but it could also rely on implicit knowledge or associative conditioning (Pavlovian associations).

The evolutionary advantage of a reward system is clear when it comes to basic (homeostatic) sensory pleasures such as taste and smell, and in social species the advantage of social pleasure in mate finding and group cohesion is also self-evident. However, it is less clear how higher-order pleasures such as artistic, monetary, altruistic, and so forth, fit in the evolutionary picture. One common view is that these awards depend on learning (the third mechanism). Some of the neural circuitry related to basic pleasure nevertheless seems to overlap.

Research on aesthetic processing of sensory perception can also shed some light on how and why a sonic environment is positively or negatively appraised. Aesthetic processing can be seen as appraisal of valence of perceived objects that comes about through a comparison between subjective awareness of current homeostatic state and exteroceptive perception of objects in the environment (Brown et al., 2011). The basic goal of this comparison circuitry is to identify whether perceived objects will satisfy or oppose our homeostatic needs. For appraisal of the sonic environment only very rarely, it could be expected that the sound object relates directly to a homeostatic need. Therefore, it is worth looking more closely at the social needs already mentioned above. The aesthetic experience of art—including music—may be argued to have social functionality, and therefore, it may have co-opted the basic circuitry used for appraisal in the context of homeostatic fulfilment. Brain imaging experiments indeed show that the same areas of the brain are activated.

Recently, Kuppens et al. (2012) reported highly ecologically valid research into the bidirectional relationship between the way we appraise our (current) environment and how that influences how we feel, plan, and act. Kuppens et al. studied this relationship in the context of core affect, which is defined as an integral blend of the dimensions displeasure–pleasure (valence) and passive–active (arousal) (Russell, 2003). Unlike emotional episodes, which are relatively infrequent, core affect is continually present to self-report.

2.5.2 Appraisal and Predictability

As already mentioned above, prediction is a logical outcome of evolution: because the context-dependent meaning of stimuli changes too frequently, evolution could not rely on instinctive responses and had to turn to associative learning mechanisms to link sensory inputs and behavioural responses. In this evolutionary framework, expected events reward the prediction system with pleasure or aesthetic emotions (Perlovsky, 2006) and a knowledge instinct is developed.

Perception of a visual or sonic environment is thus not a one-way process; the brain is constantly trying to predict the upcoming stimuli. Prediction error causes additional learning and adaptation of the prediction confidence (Winkler et al., 2009). Predictability is appraised as pleasing and aesthetic, yet too much predictability may result in a bored cognitive system. The optimal amount of predictability depends on personal characteristics and mood of the individual. Stimulus complexity and personal experience with this type of stimulus both contribute (Van de Cruys and Wagemans, 2011).

It is worth looking into music research knowledge on predictability and surprise more closely to understand the pleasure in novelty and surprise—or in other words, prediction error—and how it can affect soundscape. In Huron (2006), three kinds of response are identified: prediction, reaction, and appraisal. The prediction response has already been discussed and triggers an aesthetic emotion if the event matches expectations. The reaction response is a fast automatic response that prepares the organism for flight, fight, or freeze in case of surprise, occurring when the event does not match expectations. Finally, the appraisal response is a more leisurely process of consideration and assessment giving positive and negative outcomes. A preference for predictable events (and sounds) is explained by the anticipatory prediction success being misattributed to the stimulus itself. To explain positive emotions associated with surprise, Huron (2006) introduces emotional contrastive valence between the different expectation responses. Events that are welcome or just inoffensive but unexpected can still trigger positive appraisal. But contrastive valence also produces three kinds of pleasurable physiological response: awe, laughter, and frisson. Unexpected events (and sounds) also increase physiological arousal.

An alternative explanation for the inverted U-shape relation of aesthetics with stimulus complexity can be given by merely considering learning (Pearce and Wiggins, 2012). Extremely unpredictable stimuli afford reduced opportunities for learning, while "the learning stimulated by moderate degrees of expectation violation would be pleasurable per se" (Pearce and Wiggins, 2012, page 643).

2.5.3 Appraisal, Coping, and Behavioural Options

From a psychological perspective, appraisal of environmental stressors such as sound is often related to coping opportunities. In a primary appraisal, a person evaluates the situation with respect to its well-being. When the situation is perceived as harmful or threatening, coping resources are assessed in a secondary appraisal step. Although this appraisal and coping theory allows us to model annoyance (Botteldooren and Lercher, 2004; Maris et al., 2007), it seems to be too restricted to negative appraisal to be applicable in its basic form to the soundscape approach, except for restorative soundscapes acting as a coping resource.

The dual-phase appraisal concept can, however, be refined. For this, "coping ability" can be broadened to "opening behavioural options." If something affords behavioural options, it can be regarded as meaningful (Andringa, 2010). Sonic environments that support the behaviour that is instantaneously desired would thus be appraised positively. After a primary appraisal, behaviour options could be assessed and a secondary appraisal might follow. Sounds associated with events that open a desired behaviour option, even if unexpected, are welcomed; sounds related to events that do not prevent the behaviour option could be regarded as inoffensive.

2.6 HOLISTIC MODEL FOR SOUNDSCAPE

Classically, environmental noise has been considered a waste that needs to be prevented or mitigated in volume once the noise-producing activity has been planned. This paradigm could be related to the historical end-of-pipe approach. The soundscape approach introduces a few shifts in this paradigm. Including positive as well as negative environmental sounds in designing high-quality living environments is probably the most easily identified. However, the soundscape approach is also a user-centred approach in line with more general user-centred (product) design. The person-centred view on soundscape therefore allows understanding of many of the ideas and concepts that have been introduced by scholars and practitioners. Finally, the soundscape approach is also an integrated approach including the holistic sensory experience and all different use aspects, such as mobility, recreation, and so forth.

2.6.1 From a Use Case Perspective

In view of the user-centred approach, it is useful to discuss the holistic model for soundscape from a use case perspective. Figure 2.1 shows an example: A person has the intention to move to a place. This creates some expectations that are based on prior knowledge. When the person enters the environment, expectations are fine-tuned by observations. The intention also entails a behaviour that—in this case—is assumed to include listening for particular sounds. Observing these expected sounds leads to a pleasurable experience since it matches expectations. This affects the mood of the person and may lead to new intentions with, for example, a behaviour that

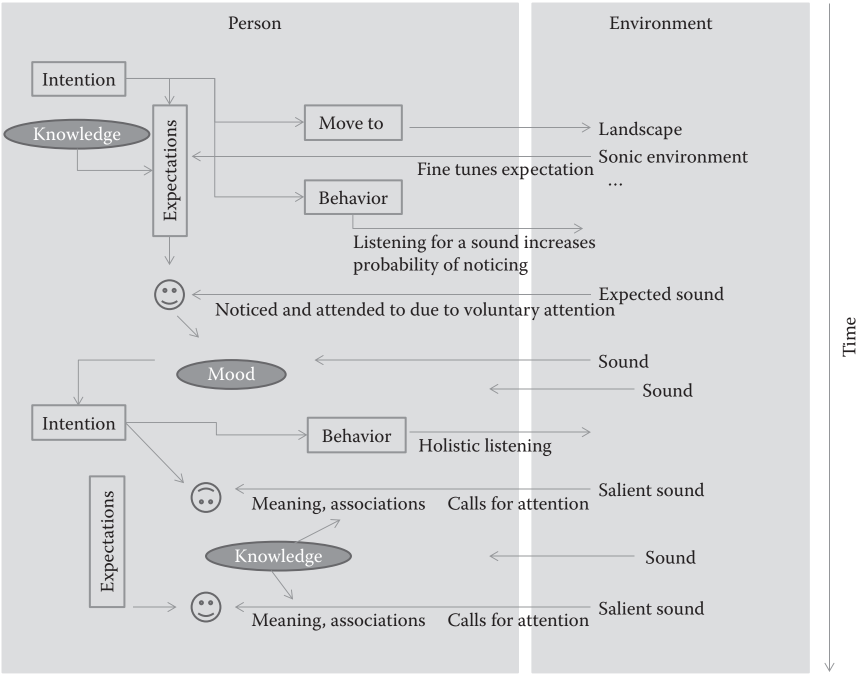

Figure 2.1 An example of a use case of a public space, focusing on environmental sound.

does not include listening in search. A salient sound may nevertheless call for attention. It could be given a meaning that makes the sound disliked. Another salient sound, although it comes as a surprise, may still generate pleasure because the meaning associated with it stimulates the intended behaviour, or at least does not offend it.

This abstract use case can be made more concrete in specific situations, yet it illustrates the complex interplay between different factors and influences.

The sequence of experiences was shown to be rather important since the listener does not respond to the current sensory input only, but also to the current sensory input interpreted and understood within a context that is created by the recent past. At the shortest time frame, meaning (e.g., sound recognition) is given within the framework of very recent experiences (e.g., sounds recently heard). Attention and inhibition of return interplay to avoid focusing on specific sensory inputs, and thus changes and transitions become much more significant than continuous audiovisual stimulation. At a somewhat longer time frame, expectations are fine-tuned and behaviour or even intentions are modified by recent experience. Transitions are therefore caused not only by changes in the environment, but also by changes in the individual and changes in location. This smoothly introduces the important role of accounting for routes and paths travelled.

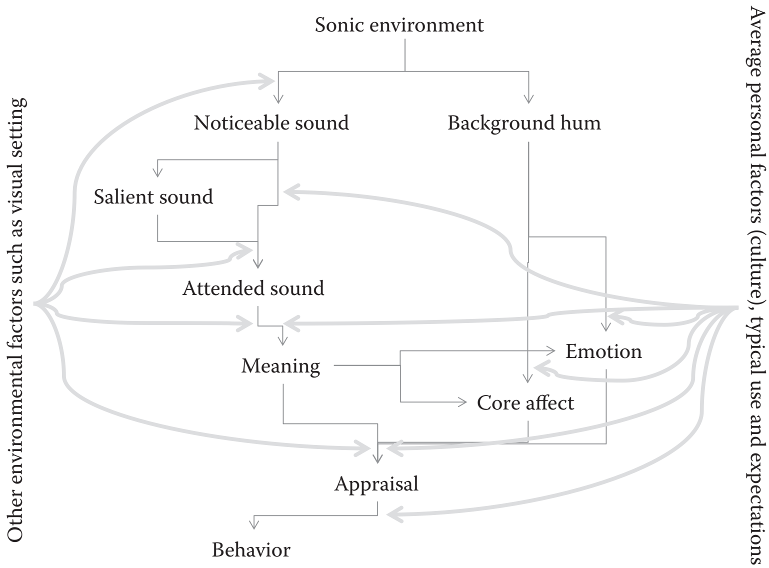

Figure 2.2 Soundscape theory outlined from the perspective of the sonic environment.

2.6.2 From the Perspective of a Sonic Environment

It is also worth looking at the conceptual holistic model for soundscape from a perspective of the sonic environment since designing a soundscape will mainly involve this sonic environment (Figure 2.2). From this perspective, personal traits, preferences, intentions, associations, and so forth, cannot be taken into account, yet the common factor between expected users of this environment can. Based on the overview given above, it can be concluded that foregrounding and backgrounding of specific sounds are important parts of listening. Foregrounding is controlled by attention, yet some sounds will not be noticeable at all, and therefore cannot be attended to. We could call this the undefined background hum. Noticeability of a sound itself depends on personal knowledge and listening capabilities, but from the perspective of the sonic environment, one can only talk about noticeability for the average listener given the environmental context, such as visibility of the sound. Whether noticeable sounds are attended to depends on the saliency of the sound, but also on many personal factors and expectations. For the latter, the typical use of the place, including typical access trajectories followed by the user of the space, is important. Noticeable sounds that do not get noticed blend into the background. Both background(ed) sound and foregrounded sounds contribute to the core affect and emotions evoked by the sonic environment. Foregrounded sounds are expected to have a stronger effect through the meaning and associations they trigger, often related to the source of the sound and the relationship of the average user to that source. Culture—regarded as a common factor in associations—and expectations will play a role. For backgrounded sound the spectrotemporal variations influence core affect and emotions directly through elements of surprise and novelty or continuity. Appraisal—mainly reported appraisal—can be interpreted as a more cognitive evaluation of the sonic environment. Both core affect and meaning influence how the sonic environment is appraised and reappraised in view of behavioural options, for example. Again, the influence of other environmental factors, culture, and expectations is significant.

It should be noted that in the model presented in Figure 2.2, the important feedback paths to modes of listening are not explicated. On an individual basis they are extremely important; however, when looking at soundscape from a perspective of the sonic environment, this iterative detail cannot be taken into account.

2.7 HOW SOUNDSCAPE THEORY CAN AFFECT PRACTICE

In this section the understanding of the processes involved in the creation of soundscape obtained above from reflection on knowledge from psychophysics, neuroscience, psychology, and so forth, is translated to practice. It gives possible directions in measuring and design. Both aspects will be elaborated on in Chapters 5 and 6 and Chapter 8.

2.7.1 Measuring Soundscape

Measuring is about representation (the justification of number assignment) and uniqueness (the representation chosen approaches being the only one possible for the object or phenomenon in question) (*Encyclopædia Britannica*). The main challenge with regard to measuring soundscape is that soundscape is a multifaceted phenomenon and hence cannot be measured with a single number.

2.7.1.1 Measuring People

When measuring people, the investigator wants minimal interference with the test persons. The observation is mainly retrospective unless subtle biomonitoring can be used. This kind of measuring can attempt to capture core affect, appraisal, restoration, and overt behaviour, and thus assesses the soundscape as a whole within a context. This type of measurement fulfils the role of creating a representation perfectly, but it is rather difficult to obtain uniqueness in the measurement. Moreover, this type of measurement should respect the way people are experiencing their environment. So the measurements should characterize not only locations one by one, but also paths between different locations.

The concept of core affect and the associated appraisal of the sonic environment appear in a number of soundscape studies. Depending on the choice of the researchers, the main appraisal dimensions are termed either pleasantness and eventfulness—which match the dimensions of core affect (Axelssön et al., 2010)—or a combination of these dimensions rotated by 45°. Cain et al. (2013) report the dimensions' vibrancy (interpreted as a combination of pleasant and eventful) and calmness (combining pleasant and uneventful). Axelssön et al. (2010) propose to interpret the vibrancy dimension as a continuum from monotonous to exciting, and the calmness dimension spanning from calm to chaotic. One could argue that these dimensions of core affect are related to the person rather than to the sonic environment, but with soundscape interpreted as an object in the mind, this does not pose any problem.

To apply this holistic approach, interviews and questionnaires are the most commonly used tool. To use the results in a planning process, information on the processes discussed above may be gathered: What sounds did people hear (attention process combined with short-term memory [Terroir et al., 2013])? What did these sounds mean to them? How does this relate to expectations concerning the place? This information should be collected after the main appraisal questions in order not to steer the attention process.

2.7.1.2 Measuring with People

Measuring with people implies that the sensory and cognitive capabilities of humans are used to assess the (sonic) environment. The participants are usually in an attentive, analytic listening mode, and thus noticeability and quality of the sound(s) per se are assessed. There is a subtle but rather important difference between this kind of measurement and the measurement of people discussed above: whether or not the sound will actually be noticed in a natural setting with persons engaged in certain activities is no longer considered. To obtain measurements that fulfil the uniqueness criterion, either one has to rely on statistical averaging of the personal factors that might influence the human observation or a master scaling (Lavandier et al., 2012) has to be used to eliminate some of these personal factors by first asking the participants to judge a set of standard stimuli. These stimuli could be either classical pink or white noise samples or reference sonic environments explicitly exhibiting the soundscape features that the research is trying to explore. The latter comes down to calibrating the human as measurement equipment.

The human observer has some capabilities that are hard to mimic using electronics and computational intelligence, for example, the capability to segregate the auditory environment into streams and objects. Thus, questions such as identifying the dominant sound source can easily be answered. Measuring with people, because of the attentive analytic listening mode, is particularly suitable for an analytic description of the soundscape.

An analytic description of the soundscape includes an inventarization of the sounds and the sources producing these sounds. It may also include a description of the quality of the sound, the meaning for that particular group of people, and an indication of congruency. The latter requires the definition of a clear context—sketched in the lab or influenced by the place in field studies—that generates particular expectations.

2.7.1.3 Measuring with Computers in a Human-Mimicking Way

Electronic equipment embodying computational intelligence can mimic the listening capabilities of humans. For the easiest indicators based on level, temporal variability of the level, spectrum, and loudness, each measurement tool produces the same unique outcome, and comparability between sonic environments becomes trivial. However, the representation of the soundscape that is created is rather poor. Information on level, spectrum, and loudness is not sufficient to allow the evaluator or designer to imagine the soundscape.

More advanced, smart sound metres are being developed that allow us to segregate the sound stream into auditory objects (Boes et al., 2012) and label these objects (Boes et al., 2013) taking into account expected sounds at a given location (Krijnders et al., 2010). Besides mimicking the auditory stream segregation, such measurement approaches also could account for the frequency of noticeability or frequency of paying attention to particular sounds (Oldoni et al., 2013). As such, these novel approaches now cover part of the measurements than can be performed with people.

Using electronic equipment has a clear advantage over measuring people: it allows for long-term monitoring. Such monitoring is necessary to study diurnal and seasonal variations in the soundscape. It is also an essential tool to detect novel and unexpected soundscape elements. However, the uniqueness requirement, which is an important factor in measurement theory, is somewhat jeopardized, as less reproducible aspects of human listening are incorporated in measurement equipment.

It should be stressed that research on measuring soundscape either with people or with human mimicking equipment is still ongoing.

2.7.2 Soundscape Design

The goal of soundscape design is to create environmental comfort by influencing the mood, the emotion, the appraisal, and the restoration of persons visiting the place. Based on the soundscape theory explained above, guidelines for future design can be formulated.

2.7.2.1 Designer's Vision and Possible Use of the Place

Modern soundscape design should start from a vision of a place and a soundscape that matches that vision. As urban design is functional design, the use of the space should be accounted for. Running typical use cases (Figure 1.1) should allow the designer to imagine and formulate the expectations of current or future users of the space. It was indeed shown that these expectations may influence the soundscape appraisal or even the mere perception of the sonic environment to a very high degree. As different uses of the space may be envisaged, careful zoning may be needed to match different expectations.

Multimodal aspects—visual, thermal, and so on—should be considered, including the spatial aspect of audiovisual matching. The path usually followed by people visiting the place has to be taken into account since recent experience has a strong influence on listening style, attention paid to sound, recognition, meaning, and appraisal.

This careful initial design phase could lead to formulating requirements for the soundscape:

- Backgrounded: Where the activity and behaviour require the soundscape to remain unattended. One should not notice that any sounds are there.
- Supportive: Where the soundscape enhances the experience and the effect of a visit to a place. The soundscape could improve the mental restoration capacity of the place (see Chapter 3); it could enhance the touristic experience, and so forth, but only as part of a multisensory experience. Specific soundmarks occasionally can attract attention or are expected as part of the experience of a place.
- Focused: Where the soundscape becomes a point of interest in itself. This can be either static, as an acoustic sculpture, or more dynamic, as surprising but pleasing sonic events, where pleasing is defined as supporting or at least not jeopardizing behaviour options.

2.7.2.2 Composition

Once the design goals have been set, the soundscape architect can start to compose the soundscape. In this, the main factors are guiding attention of the visitor and knowing the meaning of the sounds that are integrated in the composition. Attention should be purposely directed to certain components of the sonic environment, to certain sounds, while keeping attention away from unavoidable sounds that are not wanted by the designer. Sounds can be analyzed for their saliency, and for their familiarity, for the most likely users, as both determine the probability that they will be paid attention to. Visual stimuli can create opportunities for wanted sounds to attract attention, but at the same time, they can cause conflict when creating incongruence between visual and auditory stimuli. The probability that an undesired sound will attract attention anyhow, even if its source is not visible, is a crucial factor.

Composing a backgrounded soundscape is conceptually easy and largely boils down to classical noise control. Yet, taking into account saliency of the sound and congruence with the environment could allow reaching a more precise solution.

Supportive soundscapes need not only a careful selection of wanted sounds, but also a careful balance between predictability and novelty. Pleasurable elements of surprise should occur while the user is moving through the space. This allows introducing the desired amount of vibrancy or calmness, matching the whole experience. Composing a supportive soundscape is by far the most challenging.

Composing a soundscape that in itself becomes a point of interest allows for the largest amount of creativity. The auditory experience is dominated by the composition that gets the full attention of the visitor. Thus, the soundscape architect can focus on meaning, emotion, and core affect. Contrastive valence leading to awe, laughter, or frisson can be used. Examples will be given in Chapter 10.

2.8 FUTURE DIRECTIONS AND A WAY FORWARD

Understanding human (auditory) scene analysis and the important role of (auditory) attention in this process allows us to outline better assessment methods and to come to better methodologies for designing desirable soundscapes within a specific context and for a specific use. However, today the knowledge of attention mechanisms and scene analysis applicable to soundscapes has to be inferred from experimental work that uses abstract sounds in a clean context. In this chapter we attempted to do just that. The natural environment is nothing like that: it is governed by complexity, and the biological perception system has evolved to find approximate solutions for achieving goals within this complexity. Experimental research and modelling of attention and scene analysis in this natural environment have to be extended (Lewicki et al., 2014).

It has been pointed out in the previous sections that there are strong individual differences in how a sonic environment is appraised and what meaning is given to the sounds that are noticed within this sonic environment. Cultural elements and age certainly play a role and are mentioned as discriminating factors in appraisal and meaning. Most attempts to understand appraisal of a sonic environment have nevertheless focused on the average person, and indeed some general trends could be discovered from this (e.g., people like natural sounds in a park). Open-ended interviews reveal more details, but still the interviewees are likely to act as local experts assessing the environment in a pseudo-objective and rationalized way, trying to eliminate what they believe is their subjectivity. It may be worthwhile to focus more strongly on the user of a sonic environment and in particular on the diversity in individual traits, beliefs, opinions, and desires. Biomonitoring techniques could be used to assess different responses more objectively than questionnaires, even when it comes to aesthetics or pleasure.

Monitoring and simulating—as a tool for designing—soundscape requires us to account for the way a human listener perceives and understands the sonic environment within a context and use, and with emphasis on personal and cultural differences. This complex process is complicated further by the observation that the use of a place in most cases does not involve attentive listening to the sonic environment. Designing and developing computer software that can mimic this complex process and that can be used in monitoring and modelling is far from easy. Mimicking the human brain in a machine (or even a mouse brain as a starting point) is indeed identified by the European and American research funding agencies as one of the great challenges of this century. Developing machine audition can be seen as part of this challenge. Challenges that could be particularly informative for soundscape include multisensory perception, and in particular multisensory attention mechanisms; introducing learned context awareness; and modelling a biological plausible reward system, including serotonin and dopamine effects (Weng, 2013). The latter could lead to adding appraisal to machine audition in ways that are not foreseeable today.

REFERENCES

Andringa, T. C. (2010). Audition: From sound to sounds. In *Machine Audition: Principles, Algorithms and Systems*, ed. W. Wang, 80–105. IGI Global Press, 532 pages, ISBN-13: 9781615209194, August 2010.

Axelssön, O., Nilsson, M. E., and Berglund, B. (2010). A principal components model of soundscape perception. *Journal of the Acoustical Society of America*, 128, 2836–2846.

Berridge, K. C., and Kringelbach, M. L. (2008). Affective neuroscience of pleasure: Reward in humans and animals. *Psychopharmacology*, 199, 457–480.

Boes, M., Oldoni, D., De Coensel, B., and Botteldooren, D. (2013). Attention-driven auditory stream segregation using a SOM coupled with an excitatory-inhibitory ANN. In *2012 International Joint Conference on Neural Networks (IJCNN 2012)*, Brisbane, Queensland, Australia, 2012, p. 8.

Boes, M., Oldoni, D., De Coensel, B., Botteldooren, D. (2012). A biologically inspired recurrent neural network for sound source recognition incorporating auditory attention. Presented at Proceedings of IJCNN, Dallas, TX, August 4–9 2013.

Botteldooren, D., and Lercher, P. (2004). Soft-computing base analyses of the relationship between annoyance and coping with noise and odor. *Journal of the Acoustical Society of America*, 115(6), 2974–2985.

Bregman, A. S. (1994). *Auditory Scene Analysis: The Perceptual Organization of Sound*. Cambridge, MA: MIT Press.

Brown, S., Xiaoqing, G., Tisdelle, L., Eickhoff, S. B., and Liotti, M. (2011). Naturalizing aesthetics: Brain areas for aesthetic appraisal across sensory modalities. *Neuroimage*, 58(1), 250–258.

Bubic, A., von Cramon, D. I., and Schubotz, R.I. (2010). Prediction, cognition and the brain. *Frontiers in Human Neuroscience*, 4, 25.

Cabeza, R., Mazuz, Y. S., Stokes, J., Kragel, J. E., Woldorff, M. G., Ciaramelli, E., Olson, I. R., and Moscovitch, M. (2011). Overlapping parietal activity in memory and perception: Evidence for the attention to memory model. *Journal of Cognitive Neuroscience*, 23(11), 3209–3217.

Cain, R., Jennings, P., and Poxon J. (2013). The development and application of the emotional dimensions of a soundscape. *Applied Acoustics*, 74, 232–239.

Cusack, R., Decks, J., Aikman, G., and Carlyon, R. P. (2004). Effects of location, frequency region, and time course of selective attention on auditory scene analysis. *Journal of Experimental Psychology: Human Perception and Performance*, 30, 643–656.

Delaitre, P. (2013). Caracterisation des zones calmes en milieu urbain: Qu'entendez-vous par zone calme? PhD thesis, Université de Cergy-Pontoise, France.

Driver, J., and Spence, C. (1998). Attention and the crossmodal construction of space. *Trends in Cognitive Sciences*, 2(7), 254–262.

Escera, C., Alho, K., Winkler, I., and Nätänen, R. (1998). Neural mechanisms of involuntary attention to acoustic novelty and change. *Journal of Cognitive Neuroscience*, 10, 590–604.

Fritz, J. B., Elhilali, M., David, S. V., and Shamma, S. A. (2007). Auditory attention: Focusing the searchlight on sound. *Curr. Opin. Neurobiol.*, 17, 437–455.

Gaver, W. W. (1993a). What in the world do we hear? An ecological approach to auditory event perception. *Ecological Psychology*, 5(1), 1–29.

Gaver, W. W. (1993b). How do we hear in the world? Explorations in ecological acoustics. *Ecological Psychology*, 5(4), 285–313.

Genuit, K., and Fiebig, A. (2006). Psychoacoustics and its benefit for the soundscape approach. *Acta Acustica united with Acustica*, 92(6), 952–958.

Gygi, B., and Shafiro, V. (2011). The incongruency advantage for environmental sounds presented in natural auditory scenes. *Journal of Experimental Psychology—Human Perception and Performance*, 37, 551–565.

Hellman, R. P. (1982). Loudness, annoyance, and noisiness produced by single-tone-noise complexes. *The Journal of the Acoustical Society of America*, 72(1), 62–73.

Huron, D. (2006). *Sweet Anticipation: Music and the Psychology of Expectation*. Cambridge, MA: MIT Press.

ISO 1996-1. (2003). Acoustics—Description, measurement and assessment of environmental noise—Part 1: Basic quantities and assessment procedures.

Kaplan, R., and Kaplan, S. (1989). *The Experience of Nature: A Psychological Perspective*. Cambridge: Cambridge University Press.

Kayser, C., Petkov, C., Lippert, M., and Logothetis, N. K. (2005). Mechanisms for allocating auditory attention: An auditory saliency map. *Current Biology*, 15, 1943–1947.

Kirmse, U., Jacobsen, T., and Schröger, E. (2009). Familiarity affects environmental sound processing outside the focus of attention: An event-related potential study. *Clinical Neurophysiology*, 120(5), 887–896.

Krijnders, J. D., Niessen, M. E., and Andringa, T. C. (2010). Sound event recognition through expectancy-based evaluation of signal-driven hypotheses. *Pattern Recognition Letters*, 31, 1552–1559.

Kuppens, P., Champagne, D., and Tuerlinckx, F. (2012). The dynamic interplay between appraisal and core affect in daily life. *Frontiers in Psychology*, 3, 1–8.

Lavandier, C., Barbot, B., Terroir, J., and Schuette, M. (2012). Calibration of subjects with master scaling: An application to the perceived activity disturbance due to aircraft noise. *Applied Acoustics*, 73, 66–71.

Lavie, N. (2010). Attention, distraction, and cognitive control under load. *Current Directions in Psychological Science*, 19, 143–148.

Lavie, N., and Tsal, Y. (1994). Perceptual load as a major determinant of the locus of selection in visual attention. *Perception and Psychophysics*, 56, 183–197.

Lewicki, M. S., Olshausen, B. A., Surlykke, A., and Moss, C. F. (2014). Scene analysis in the natural environment. *Frontiers in Psychology*, 5, 199. doi: 10.3389/fpsyg.2014.00199.

Maris, E., Stallen, P. J., Vermunt, R., and Steensma, H. (2007). Noise within the social context: Annoyance reduction through fair procedures. *Journal of the Acoustical Society of America*, 121, 2000–2010.

Mayr, S., Buchner, A., Moller, M., and Hauke, R. (2011). Spatial and identity negative priming in audition: Evidence of feature binding in auditory spatial memory. *Attention Perception and Psychophysics*, 73, 1710–1732.

McAdams, S. (1993). Recognition of sound sources and events. In *Thinking in Sound: The Cognitive Psychology of Human Audition*, ed. S. McAdams and E. Bigand. Oxford: Clarendon Press.

Mondor, T. A., Breau, L. M., and Milliken, B. (1998). Inhibitory processes in auditory selective attention: Evidence of location-based and frequency-based inhibition of return. *Perception and Psychophysics*, 60, 296–302.

Niessen, M. E., van Maanen, L., and Andringa T. C. (2008). Disambiguating sounds through context. In *2008 Second IEEE International Conference on Semantic Computing (ICSC)*, 2008, pp. 88–95.

Oldoni, D., De Coensel, B., Boes, M., Rademaker, M., De Baets, B., Van Renterghem, T., and Botteldooren, D. (2013). A computational model of auditory attention for use in soundscape research. *Journal of the Acoustical Society of America*, 134, 852–861.

Payne, S. R. (2013). The production of a perceived restorativeness soundscape scale. *Applied Acoustics*, 74, 255–263.

Pearce, M.T., and Wiggins, G.A. (2012). Auditory expectation: The information dynamics of music perception and cognition. *Topics in Cognitive Science*, 4, 625–652.

Perlovsky, L. I. (2006). Toward physics of the mind: Concepts, emotions, consciousness, and symbols. *Physics of Life Reviews*, 3, 23–55.

Prime, D. J., Tata, M. S., and Ward, L. M. (2003). Event-related potential evidence for attentional inhibition of return in audition. *Neuroreport*, 14, 393–397.

Russell, J. (2003). Core affect and the psychological construction of emotion. *Psychological Review*, 110(1), 145–172.

Sandrock, S., Griefahn, B., Kaczmarek, T., Hafke, H., Preis, A., and Gjestland, T. (2008). Experimental studies on annoyance caused by noises from trams and buses. *Journal of Sound and Vibration*, 313(3), 908–919.

Santangelo, V., Fagioli, S., and Macaluso, E. (2010). The costs of monitoring simultaneously two sensory modalities decrease when dividing attention in space. *Neuroimage*, 49(3), 2717–2727.

Shamma, S. A., Elhilali, M., and Micheyl, C. (2011). Temporal coherence and attention in auditory scene analysis. *Trends in Neuroscience*, 34, 114–123.

Snyder, J. S., and Alain, C. (2007). Sequential auditory scene analysis is preserved in normal aging adults. *Cerebral Cortex*, 17(3), 501–512. doi: 10.1093/cercor/bhj175.

Talsma, D., Senkowski, D., Soto-Faraco, S., Woldorff, M. G. (2010). The multifaceted interplay between attention and multisensory integration. *Trends in Cognitive Sciences*, 14(9), 400–410. doi: 10.1016/j.tics.2010.06.008.

Terroir, J., De Coensel, B., Botteldooren, D., and Lavandier, C. (2013). Activity interference caused by traffic noise: Experimental determination and modeling of the number of noticed sound events. *Acta Acustica united with Acustica*, 99(3), 389–398.

Truax, B. (2001). *Acoustic Communication*. 2nd ed. Westport, CT: Ablex.

Van de Cruys, S., and Wagemans, J. (2011). Putting reward in art: A tentative prediction error account of visual art. *Iperception*, 2, 1035–1062.

Weng, J. (2013). How the brain-mind works: A two-page introduction to a theory. *Brain-Mind Magazine*, 2(2).

Winkler, I., Denham, S. L., and Nelken, I. (2009). Modeling the auditory scene: Predictive regularity representations and perceptual objects. *Trends in Cognitive Sciences*, 13, 532–540.