Same sound – Different meanings: A Novel Scheme for Modes of Listening

Kai Tuuri, Manne-Sakari Mustonen, Antti Pirhonen
Department of Computer Science and Information Systems,
P.o. Box 35, FI-40014 University of Jyväskylä, Finland
{krtuuri, msmuston, pianta}@cc.jyu.fi

Abstract. This paper is grounded on the *multimodal listening* hypothesis, which suggests that listening is a multi-focused process based on multiple distinct, environmentally shaped activating systems and listening strategies. Different modes of listening can operate concurrently complementing each other with different perceptual perspectives. In sound design, the potential of this hypothesis lies in its ability to account for the heterogeneity of different, even contradictory levels of interpretations, meanings and emotions evoked by the same listening experience. In this paper the theoretical basis of listening modes is further analysed and reflected upon in the context of sound design. We propose a comprehensive scheme of eight modes of listening (reflexive, connotative, causal, empathetic, functional, semantic, critical, and reduced) accompanied by examples of their significance in sound design for user-interfaces.

1. Introduction

In contrast to hearing, listening is an active process that provides a means to pick out information for our needs from the auditory environment. It is usually associated with voluntary attention and focusing on something. At present, a reasonable amount of studies exist concerning the area of auditory perception [see 1]. However, psychoacoustic models in which perception is built up from low-level perceptual atoms are inadequate for understanding creation of meanings. In most cases we do not perceive sounds as abstract qualities; rather, we denote sound sources and events taking place in a particular environment (e.g. dog barking) or we concentrate on some other level of information. Apart from a few examples in previous literature, modes of listening have received surprisingly little attention.

Listening is highly multimodal activity in nature. Multimodality of listening means that there are several distinct strategies to listen. It is a distinction of listening strategies and perceived experiences, not a distinction between sounds - although some sounds encourage the use of certain modes more strongly than others. Each mode of listening considers its own source of information in the auditory stimuli either with or without its context. The same sound can essentially be listened to with different kinds of attention and with different outcomes. Despite their separate nature, modes can and often do operate concurrently complementing and influencing each other.

How many ways can we listen to the same single sound? The process of listening to speech provides an example of a variety of perceived meanings. In addition to the conventional (linguistic) meaning of speech we can focus on listening to a speaker as a sound source (e.g. gender, age, dialect, emotional state). Or we can concentrate on the interactive nature of speech in a conversational context (e.g. getting attention, approving, encouraging). We can also pay attention to the qualities of speaker's voice (e.g. timbre, melodic contour, rhythm). Such meanings can even be perceived as contradictory e.g. when non-verbal cues in speech do not match with the verbal content.

The exploration of the multimodality of listening experiences promotes a better understanding of how meanings can be conveyed in effective sound design. This is our primary motivation to study this subject. In the following text previous accounts of the modes of listening are reviewed and then a revised account is proposed and detailed with examples.

2. Previous accounts of listening modes

2.1. Everyday listening

Our everyday listening is not focused on sounds. Instead, we usually hear sound sources: actions and events that cause sounds. We hear footsteps on a sidewalk, a car passing by, breaking of glass etc. We might also try to figure out how far and in what circumstances these events happen as we use listening to outline our environment to support our actions. This source-orientated mode of listening seems to be so effortless that we are not conscious of it. Listening studies by Vanderveer, Ballas and Gaver [2] show evidence that subjects indeed tend to describe a sound by its source or an event that caused it. In the case of ambiguous sounds, confusion and misidentifications are argued to be based on similarities in the mechanical structures of source events (such as "hammering" and "walking"). From the perspective of ecological (gibsonian) perception such confusions relate to shared properties in the affordance structures of various sound events. [3] In such cases, additional contextual information is required to confirm the denotation of the sound source.

The automatic nature of source-orientated listening and the phenomenon of ambiguity is frequently exploited in *Foley-tradition* of sound design [4] by framing non-authentic but believable (i.e. affordable) sounds persuasively into a suitable narrative context. As a part of the craftsmanship of a sound designer is required to identify essential components of sound which can convey a desired narrative effect (e.g. denotation of an event in a fictional environment). The freedom to use non-authentic sound sources gives a designer a much wider range of possibilities to suitably enhance an audience's experience. In contemporary audiovisual narration, even source-visualised on-screen sounds are often produced or treated artificially.

2.2. Reduced and musical listening

Possibly the earliest explicit mention of the modes of listening in previous literature can be found in the work of Pierre Schaeffer [5]. He proposes a distinguished *reduced* mode of listening by which we intentionally divorce the phenomena of sound from any functions as a medium of signification. An objective listening perspective was created to manage and handle sounds as abstract and fixed objects (*objet sonore*) for composition purposes of the *musique concréte* tradition. In

Schaeffer's phenomenology, reduced listening is a mode where the sound (and its qualities) is perceived *per se* resisting any claims about the exterior world. Such a totally abstract and "meaningless" perception of sound is of course a purely theoretical concept, and Schaeffer's work has received criticism as such [6].

Gaver's distinction of *musical listening* as opposed to everyday listening [2] shares essential similarities with the definition of reduced listening. Schaeffer's thoughts highlighted the fact that sounds from the everyday world can be used and listened to musically. The mode of musical listening is not intended to be restricted to sounds defined as music. Gaver thus recognises the "cross-referential" nature of different modes; it is possible to listen everyday auditory environments as music (e.g. attending to pitch contours and rhythm) and conversely possible to listen to musical performances in terms of causality and sound sources (e.g. separating different instruments or stems). Therefore as a term, musical listening can be misleading because music can be listened to in various modes – not just as abstract structures. Reduced listening can be referred to as musical listening only when listening process is concerned with musically determined qualities and structures.

However, the reduced mode of listening was later applied to film sound design by Chion [7]. Here the listening experience was objectified, by voluntarily resisting the natural denotation of a sound source or its meaning. By concentrating on the sound itself, sound designers could "open their ears" to the abstract qualities of sound. In this way, more creative or effective ways to utilise sound in narrative or artistic context became possible. The idea of reduced listening stresses the (analytic) perspective in which it is more important to understand what the sound *sounds* like than how it has been produced (see e.g. Foley tradition). Unlike Schaeffer, Chion puts forward the idea that reduced listening is more of a tool for analytic discovery of sound beyond its evident denotative meaning.

Both Schaeffer's and Chion's views are concerned with *acousmatic* situations [5,6,7] where the actual cause of sound is hidden from listener. This is indeed the case with mechanically reproduced and transmitted sounds e.g. telephone, radio and recordings. Acousmatic sounds thus permit more freedom of imagination for a listener to form a sound-only based perception and allow sounds to be composed artificially often by combining a variety of natural or non-natural sources. Acousmatic situations however do not automatically encourage reduced listening. Chion suggests that they can even intensify the motivation for everyday listening (i.e. causal listening, see 2.3.) when the visual support is removed [7].

2.3. Three listening modes of Chion

In his book Audio-Vision, Michel Chion was the first to introduce a more comprehensive scheme for modes of listening [7]. It consists of *causal*, *semantic* and *reduced* modes (see Figure 1). Two of them, causal listening (i.e. everyday listening) and reduced listening, were discussed above. The third mode, semantic listening, focuses on conventional meanings that the sound might represent by code, language or habit. As causal (everyday) mode of listening refers to ecologically-orientated evident[1] denotations, there was indeed a place for a mode that addresses socio-culturally shaped and learned (symbolic) codes. Spoken language is the most obvious example of cultural convention, but semantic listening can refer to anything which creates a meaning for a sound that is not "literally"[2] there. Examples could range from exact rule-type codes (e.g. Morse code) and natural languages to more passively learned pragmatic habits (e.g. an applause after a performance) or conditioned associations (e.g. an ambulance siren). Many codes are so deep-rooted in the form of dispositions or habits, that they appear almost innate to us.

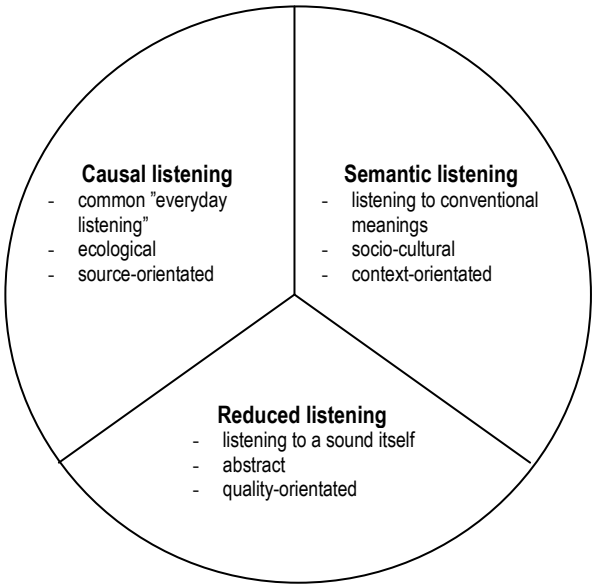

Figure 1. Three listening modes formed on the basis of Chion's classification [7]

Chion's scheme of listening modes is quite well known and it has proven its usefulness for sound designers. It is comprehensive and has broad categories that are easy to understand. However, it fails to capture more refined distinctions between clearly separate modes of listening, which will be covered later in this paper in the form of a revised model. Nevertheless, Chion's scheme forms the basis for our development of a more detailed scheme.

2.4. Activating systems

David Huron has suggested a six-component theory of auditory-evoked emotion. Despite his primary interests in emotions, his theory forms a relevant perspective to the modes of listening. Huron assumes a bio-cultural perspective on emotions as adaptive behaviours expressed within and shaped by an environment. Six *activating systems* are determined (see Table 1). These are evolved to serve specific functions and they all are capable of operating concurrently and evoking various emotional states. [9]

We assume that activating systems are not restricted to evoking emotions only, but in a similar way they can evoke other kind of meanings. Besides the denotative system which is linked to causal listening, the rest of the activating systems concerns novel and complementary perspectives to listening modes.

[1] Evident meanings, as defined by C.S. Peirce, are referred to as *iconic* and *indexical* relations of a sign and the object it refers to [8].

[2] By literal meanings we refer to ecologically inferred denotations, such as "meow" means a cat object.

Table 1. Brief summary of activating systems [9]

Reflexive system	Fast physiological responses
Denotative system	Processes which allow listener to identify sound sources
Connotative system	Processes that allow listener to infer various physical properties and passively learned associations (e.g. from temporal patterns)
Associative system	Arbitrary learned associations
Empathetic system	Allows the listener to perceive cues that signals someone's state of mind (an agent causing the sound)
Critical system	Reflective self-monitoring concerning the verification of perception and the appropriateness of one's responses

2.5. Functional perspective

Every sound which is intentionally used for some purpose has a specific function. When a listener answers a question "what is the purpose of that sound?" she defines the perceived function of the sound that is used in a particular context. We might become aware of the functions e.g. when the sound is perceived as a fire alarm, or when we feel music suitable for relaxing purposes, or when we perceive a sound effect as a transitional cue in audiovisual narration. The functional perspective of sound was explored in previous literature e.g. by Hermann and Ritter [10], and Jørgensen [11] who has examined the functional aspects of game audio. Roman Jakobson's model of communicative functions [12] is also related to this perspective.

Sound as a function can be seen as a pragmatic frame for meaning. The way that sound appears in a functional context affords a certain perspective to the process of interpreting the meaning of sound. The procedural chain of events, actions and causalities in a situation can give an indicative meaning even to a "meaningless" beep.

Although the sound itself can suggest the purpose of its use, perceiving the function of sound requires an awareness of context. The context concerns equally the situational factors as well as the listener's past experience on similar functionalities. The perception of a function is often related to a certain framework of common habits (e.g. habits of non-verbal interaction, conversation, musical performance, audiovisual narration or different user interfaces). The function of sound is particularly important in practices of interactive communication.

Functional semantics of sound can be seen as a distinct level of meaning which indicates and/or promotes a functional purpose of sound. It is indicated by the situational context but can also be indicated by the sound. For an example; let's consider the spoken word "dad", whose "pure" verbal meaning of word is carried by the sound. If we shift our attention to the source of sound we can tell that it is a voice of a child. As we concentrate more on the voice and the way the word was spoken (prosodic qualities of voice), we can guess that a child is demanding attention from her dad. She is definitely not just mentioning the word "dad". The attention-demanding function of this utterance can even be perceived without understanding the verbal content.

3. Hierarchical account of listening modes

Previous accounts of listening modes have been incoherent and limited in their scope Therefore, we have formulated a new scheme. We have applied various relevant perspectives in order to form a more detailed and comprehensive outline of the listening modes. It is intended that this new outline will be utilised by audio designers and tested by audio researchers.

The basis of our pursuit of new categories is Chion's determination of three modes. One of the obvious shortcomings of the original scheme was its inability to consider connotations of sound. For the sake of interactive communication, we believe that perspective of functions of sounds deserves attention. Furthermore, activating systems introduced by Huron [9] offers relevant perspectives for developing a new scheme.

We propose a novel, *hierarchical* scheme of modes of listening (see Table 2) which consists of two pre-conscious modes (*reflexive* and *connotative*), two source-orientated modes (*causal* and *empathetic*), three context-orientated modes (*functional*, *semantic* and *critical*) and a reduced mode. The order of modes implicates their level of cognitive abstraction from low to high.

3.1. Reflexive mode of listening

In reflexive listening the focus is on automatic and fast audio evoked responses. Huron mentions various reflexes [9] including orientating response, startle response, defence reflex and reflexive responses that relates to expectations, habituation, sensory dissonance and attention. These responses resist conscious mediation, so in a strict sense this category cannot be considered as a pure mode of listening as it is impossible to control or focus on automatic reflexes themselves. In any case, reflexive responses represent clearly an important way by which meanings and emotions can be evoked by sounds.

3.2. Connotative mode of listening

In connotative listening, the focus is on early associations that the process of listening pre-consciously evokes. These associations are references made by similarity to past experiences of a listener, prior to any reasoned denotations. A French semiotician Roland Barthes [13] has concluded: "…denotation is not the first meaning, but pretends to be so…". He implies that even though it appears that we can reason a denotative meaning instantly, it is only an illusion because we have already made a number of connotative associations. Denotation can then be defined as a "final", more reasoned connotation. The important thing about connotation to realise is that besides denotative meaning, various connotative meanings can arise from the arsenal of excess associative "building material".

At the most primitive level, connotative processes permits a listener to infer various physical properties of sound. These properties can indicate perceptual information concerning the sound source and environment: size, material, energy, proximity and excitation. Connotations can also be evoked from certain mechanical structures of source events (e.g. temporal patterns that evokes a "galloping-like" meaning) even if the sound source has nothing to do with what it connotates. [9] Besides the physical and ecological environment, associative cues can also relate to arbitrarily learned cultural experiences. For that reason we propose that the connotative mode of listening is not only linked to processes of connotative but also to processes of associative activating system.

Like the reflexive mode, the connotative mode of listening is involuntary in nature. Therefore focussing on connotations can be somewhat established by mentally exploring free associations and voluntarily resisting denotations.

Table 2. Summary of the revised scheme of listening modes with examples

Type:	**Mode:**	**Questions:**	**Example:** *Loud sound of train whistle (in a movie)*
Pre-conscious modes:	**Reflexive**	Did you notice any reflexive responses triggered by sound?	Startle response!! It alarms and grabs an attention.
	Connotative	Can you describe what kind of freely formed associations listening immediately evoked?	Big...strong...lots of power...close proximity ...screeching...air blowing..whistle..scream... ....old steam trains....western movies....
Source-orientated modes:	**Causal**	What could have caused the sound?	It´s a train. *Critical second thought*: the sound comes from the TV.
	Empathetic	Does it feel that sound signals someone's state of mind or intentions?	Whistle sounds feels desolate and angry.
Context-orientated modes:	**Functional**	What was the purpose of the sound? What function does the context indicate?	The driver signals train's departure. *Critical second thought*; sound is used as transitional cue between scenes (just before a visual cut to railway station).
	Semantic	Does the sound seem to represent any symbolic/ conventional meanings?	The whistle represents pain... of a suffering man (by replacing his scream)
	Critical	Was the sound suitable for the situation? Did you understand it correctly?	Ah, no panic. That sound belongs to the movie. It was a cliché but quite effective.
Quality-orientated mode:	**Reduced**	Can you describe the properties of the sound itself as objectively as possible?	Sound is high-pitched and loud. A big contrast against quiet earlier scene.

3.3. Causal mode of listening

In causal listening the focus is on denotation of the source of sound and determination of an event that caused the perceived sound. This mode of listening is derived from the scheme of Chion (see 2.1 and 2.3.). This mode is also directly linked to denotative activating system. Causal listening is often referred to as a mode of common everyday listening.

3.4. Empathetic mode of listening

In empathetic listening the focus is on cues that could signal someone's state of mind. Empathetic mode of listening is thus directly linked to the empathetic activating system [9]. It is closely related to causal listening, in the sense of considering the possibility of a human or animal as a sound source or cause of sound. It is also related to connotative listening in the sense of potential auditory evoked associations (e.g. from intensity or certain rhythmic pattern) which can refer to emotional states, intentions or even communicative functions. For example listener can recognise a sad or nervous voice. On the other hand, listener can perceive e.g. a loud slamming (lots of energy) of a door as a possible expression of anger.

3.5. Functional mode of listening

In functional listening the focus is on the purpose of a sound in its context (see 2.5.). This mode considers the possibility that the sound is used for some specific function, which is pragmatically indicated by a sound in relation to the context. In the domain of non-speech auditory cues, a perceived function can be e.g. attention-demanding, alarming, orientating, approving, prohibiting, marking, prompting, giving a feedback or noticing.

3.6. Semantic mode of listening

In semantic listening the focus is on denoting any conventional meanings that a sound might represent. This is the second mode of listening which is derived from the scheme of Chion (see 2.3.). By semantic listening lower-level meanings are also reorganised for conventional reasoning to take cultural context (habits, codes) into account.

3.7. Critical mode of listening

In critical listening the focus is on the reflective judgement of auditory perception. As a mode of listening, it applies the idea of Huron's [9] critical activating system. Critical listening concerns appropriateness or authenticity of sound in a given context. It also considers the appropriateness of one's responses. That includes judgements of possible misunderstanding, deception, false urgency or generally the need to be concerned with the sound. Additionally, at its highest level, critical judgements can be based on aesthetical dispositions.

3.8. Reduced mode of listening

In reduced listening the focus is on the sound itself and its qualities. This is a third mode of listening which is derived from the scheme of Chion (see 2.2. and 2.3.). The examination of sound phenomena itself requires that a listener is consciously resisting any denotations of a sound source or its meaning. This mode of listening is thus exceptionally voluntary and very likely requires high-level cognitive abstraction.

4. Observations from a case study

For a few empirical observations of modes of listening, we present some examples from a group panel discussions of our earlier case study (see Pirhonen et al. [14] which addresses the development of the design panel methodology). Although the panels were not originally conducted for the study of the listening modes, and during the panel sessions the listening modes were not considered, some examples can be pointed out. In the study, the designing of user interface sounds were studied in a series of group panel discussions. A group of panellists carried out design tasks in an iterative fashion; first for idea generation and then for evaluation of the designed outcome. The target of sound design was the user interface of a physical browsing application [15] used in a bicycle exhibition. Each panel session had different tasks and goals. The examples, relevant to this study, consider a warming-up task from the first session and sound evaluating tasks from the third panel session.

Before the actual design tasks, which was the purpose of our panel sessions, the moderator played some soundscape samples to the panellists and after that panellists discussed what they heard. These soundscape tasks were conducted to "open the ears" of the panellists.

The first soundscape sample was recorded in a bird watch tower during morning hours. The panellists described the

environment by the objects causing sounds, such as birds, wind humming in the trees, distant road humming (causal listening) and by connotations and by descriptions such as "trees sound as green as in our summer cottage in the relaxed summer morning, birds sound happy" (connotative and empathetic listening). The sounds of the distant motorway were considered as not suitable to the otherwise relaxed atmosphere (critical listening).

The other listening sample was little walkthrough from an elevator to a silent entrance hall, and from there into a noisy rush hour restaurant. Panellists described the entering from silent room to the noisy restaurant as a defence-reaction evoking event; the sound-mass of the restaurant was described as angry, scary and stressing rush-noise (reflexive, connotative and empathetic listening). The noisy environment was also described as unapproachable and unpleasant (critical listening). The moderator told the panellists before the sample that in the beginning there is a bit noise due to the recording technique. That was actually not true, the recording was clean. We observed that this additional task orientation of considering some inappropriate sounds affected the listening experience of the panellists. On this second task they were generally more critical and analytical listeners and described e.g. the humming and clicking sounds (of the elevator) and considered their appropriateness, whether the sound was original or an error of the recording.

After the ear-opening tasks, the group started the actual design tasks for that panel session. First the panellists familiarised themselves with the application by listening to a use scenario in the form of radio-play narration. Actual events-to-be-sonified (successful activating of a physical link, process of loading, loading ready) were clearly indicated in the story. Secondly, candidate sounds were played alone sequentially for each function and the most promising sounds were voted to go through to the next phase. The third phase of the panel session was that the sounds were played within the radio play, so that the panellists obtained a more holistic experience of the use situation, and heard the nominated UI-sounds connected to the procedure of the use scenario.

When the sounds were played alone, they were mostly judged by the criteria of subjective satisfaction and connotations like "this sound is not good, I do not like it, it sounds too lazy" etc. or "this is good, happy sound, I like it" or "it was good, snappy and attention grabbing sound, suits for the function". Some sounds were voted to go through to the next phases, others were rejected. One example of the rejected sounds was rattling-like sound which was designed to indicate the loading process. When the sound was judged alone, it was considered as irritating clacking of teeth when feeling cold, and one panellist said that it reminded her of an annoying little boy with a ratchet (connotative, causal, empathetic and critical listening). The sound was rejected as too annoying and not suitable for the context.

In the next phase the panellists heard the radio play again, this time with the selected UI-sounds played within the story. Now, as the panellists were more immersed in the functional context and able to experience the whole situation procedurally, some of the earlier judgements changed. Some sounds that were judged as effective in the second phase were no longer considered suitable for the context, and some rejected sounds were asked to be elevated again. A more important factor than the subjective effect of the sounds was the match with sonic functions and events. During this last phase, the rejected rattling sound was now judged as the top-rated sound as an indicator for the loading process. Now the same sound was described as sound of small cogwheels, indicating the function of something happening, rather than the annoying ratchet or teeth chattering. This example indicates how the functional context provides a crucial part of understanding the sound.

5. Listening modes and the current design paradigms of user interface sounds

In the research field of auditory cues in user interfaces (UI) there has been only little discussion concerning the multi-faceted meaning construction described in this paper. In 1994, the workshop report of CHI'94 discusses that "A more central concern was how to effectively convey information with non-speech auditory cues. Sounds can be interpreted at several levels…Current user interfaces have not yet addressed this in deeper expressive level in their use of sounds." [16]. Despite the early recognition of the problem, it has not been widely considered within the research paradigm of UI-sounds.

In this paper we proposed a novel scheme for the modes of listening, which comprehensively binds together somewhat scattered discussions of earlier literature that concerns the issue. Listening modes play an important role as a tool for sound design. As we can listen to a sound with various forms of attention, the process of sound design must then concern auditory signs from various perceptual perspectives, in order to ensure consistent support for common communicative goals. The worst case design scenario would be that different listening modes evoke contradictory meanings (e.g. function implies an alarm sound, but the major chord invokes positive or happy connotations), or when the sound is experienced as annoying despite its perceived relevance in the context.

To demonstrate the relevance between sound design and perspectives of different listening modes, let's first examine two seemingly opposite design paradigms of user interface sounds: *earcons* and *auditory icons*. Earcons are defined "…as abstract, synthetic tones that can be used in structured combinations to create sound messages to represent parts of an interface." [17]. The symbolic relationship between the sound and its meaning is seen beneficial as sounds do not have to correspond to what they represent [18]. Meanings are thus arbitrarily coded and therefore learning of specific codes is required prior to effective understanding. It seems that the philosophy behind current earcon design is related to information theory [19] with implicit assumption of the role of sound as a carrier of coded information. Current earcon design guidelines consider sound by emphasising psychoacoustic phenomenon on how sound may be masked or how sound stream can be segregated (judgements on timbre, register, rhythm, concurrent sounds etc.) [20]. These channel-orientated perspectives and considerations of channel-noise factors (e.g. masking) further emphasises the information theory based view of communication.

The paradigm of auditory icons, conversely, relies on iconicity and the ecological perspective of auditory perception [21, 3]. This essentially means that when listening we naturally pick up recognisable parts from the auditory stimuli. Relations between sound and its meaning are based on similarities with familiar aspects of our everyday environment. They can be denotations of sound sources or partial indicators that point to some mechanical properties of a sound causing event. In sound design, similarities can also be used in a metaphorical way. The most important difference to the earcon-paradigm is that the design of auditory icons is more focused on how the sound itself, by resemblances, motivates the meaning-creation. Just as in the traditional film sound design, meanings appear to be motivated by the sound.

We can conclude that the earcon-paradigm is concentrated mainly on two modal perspectives: semantic mode (extreme requirement of code) and reduced mode (sound is supposed to be heard as musical parameters). The design of alarm sounds

additionally concerns reflexive mode of listening. On the other hand the auditory icon paradigm is determined on perspectives of causal mode (source recognising), connotative mode (physical property indicators of an event) and in some sense also functional mode (meanings of sounds shares iconic similarities with the event it represents in application). We thus find that the distinction between earcons and auditory icons is not intended as a distinction between UI-sounds themselves. In fact, that categorisation seems to be more related to which modes of listening are adopted for the paradigm in question. In light of listening modes, earcons and auditory icons are to be considered as *design paradigms* – not as necessary distinct types of UI-sounds. The cross-related nature of listening modes allows for the consideration of different design paradigms in tandem.

Indeed, an optimally designed earcon can also utilise its expression with e.g. iconic and affective levels of meaning with cues to some familiar qualities or habits of the experienced world - even when an abstract form of expression is chosen.

6. Conclusions

We feel that our own main contribution of this paper is the systematic review and synthesis of listening modes, and within the proposed scheme the inclusion of an explicit functional mode of listening. We argue that the purpose of sound in a functional context is an important factor in *interaction design* within user interfaces. Firstly, every user interface design must address the role of a sound in interaction. Secondly, the perceived function of sound in a situational context represents itself as an important class of meanings. A user can get context-derived indications to suitably interpret even a "meaningless" beep, not to mention sounds that convey some additional semantic support for the appropriate interpretation of meaning. We can find that the two classic design paradigms discussed above (earcons & auditory icons) are deterministically more concerned with the semantic aspects of UI-sound element itself – not the aspects of how the sound is used in the functional context of UI. The complementary perspective we propose is more procedural in nature; in this approach it is more important *how* meaning is created in a given context than *what* the meaning is *per se*. Functional semantics of sound is based on tacit reasoning and pragmatically evoked semantics.

The general process of sound design for a user interface, at least implicitly, should begin with exploring the relevant communicative purposes of sound in UI-interaction. The outcome of that meta-design process will be a list of functions of sounds referring to various events and processes taking place when user tasks are performed. By analysing those functions within those scenarios and situations, a designer can find associative ideas for relevant functional semantics for the actual sound design.

The scheme for modes of listening, which is presented in this paper, is intended to open a discussion concerning the topic. Also this study is to be seen as compiling a review of various aspects concerning the complex scheme of meanings inferred from sound. The new perspective, the functional mode of listening, is the most prominent contribution from the perspective of audio interaction design and research. The observations from our case study support our assumptions. Nevertheless, more empirical evidence should be gathered to validate the appropriateness of modes of listening for user interface design. Sound design cannot afford to overlook the diversity of meanings and the affective responses that the sound evokes in the context of its use. As a conceptual model, the proposed scheme of modes of listening can guide the designer to find answers to that challenge.

Acknowledgements

We thank Professor Tuomas Eerola for his support and encouragement for our work. Many thanks to Emma Murphy for her comments and help in writing English.

This work is funded by Finnish Funding Agency for Technology and Innovation, and the following partners: Nokia Ltd., GE Healthcare Finland Ltd., Sunit Ltd., Suunto Ltd., and Tampere City Council.

References

[1] Bregman, A. Auditory Scene Analysis. Cambridge, MA: MIT Press (1990)

[2] Gaver, W. The SonicFinder: An Interface That Uses Auditory Icons. Human-Computer Interaction 4, 1 (1989), 67-94

[3] Casey, M. Auditory Group Theory: with Applications to Statistical Basis Methods for Structured Audio, Ph.D. Thesis, MIT Media Lab (1998)

[4] Mott, R. Sound Effects: radio, TV, and film. Boston, MA: Focal press (1990)

[5] Schaeffer, P. Traité des objets musicaux. Paris: Editions du Seuil (1968)

[6] Kane, B. L'Objet Sonore Maintenant: Pierre Schaeffer, sound objects and the phenomenological reduction. Organised Sound 12, 1 (2007), 15–24

[7] Chion, M. Audio-Vision: sound on screen. New York, NY: Columbia University press (1994)

[8] Peirce, C.S. What is a sign? In Essential Peirce: Selected philosophical writings vol. 2. Bloomington, IN: Indiana University Press (1998), 4-10

[9] Huron, D. A six-component theory of auditory-evoked emotion. In proceedings of ICMPC7 (2002), 673-676

[10] Hermann, T & Ritter, H. Sound and Meaning in Auditory Data Display. In proceedings of the IEEE, vol. 92, 4 (2004), 730-741

[11] Jørgensen, K. On the Functional Aspects of Computer Game Audio. In proceedings of Audio Mostly 2006, 48-52

[12] Jakobson, R. Closing Statements: Linguistics and Poetics. In Sebeok, T. A. (ed.) Style In Language, Cambridge, MA: MIT Press (1960), 350-377

[13] Barthes, R.. S/Z. New York: Hill & Wang (1974)

[14] Pirhonen A, Tuuri, K., Mustonen, M-S. & Murphy, E. Beyond Clicks and Beeps: In Pursuit of an Effective Sound Design Methodology. In proceedings of HAID2007 (in press)

[15] Välkkynen, P. Hovering: Visualising RFID hyperlinks in a mobile phone. In proceedings of MIRW 2006, 27–29

[16] Arons, B. & Mynatt, E. The future of speech and audio in the interface. In SIGCHI Bulletin 26, 4 (1994), 44-48

[17] Brewster, S, Wright, P. & Edwards, A. A detailed investigation into the effectiveness of earcons. In Kramer, G. (ed.) Auditory display. Reading, MA: Addison-Wesley (1994), 471-498

[18] Blattner, M, Papp, A. & Glinert, E. Sonic Enhancement of Two-Dimensional Graphic Displays. In Kramer, G. (ed.) Auditory display. Reading, MA: Addison-Wesley (1994), 447-470.

[19] Shannon, C. & Weaver, W. The Mathematical Theory of Communication. Urbana, ILL: University of Illinois Press (1949)

[20] Brewster, S., Wright, P & Edwards, A. Experimentally derived guidelines for the creation of earcons. In Adjunct Proceedings of HCI'95, 155-159

[21] Gaver, W. Auditory Icons: Using sound in computer interfaces. Human-Computer Interaction, 2. (1986), 167-177