T. Ziemer, "Sound Terminology in Sonification,"
J. Audio Eng. Soc., vol. 72, no. 5, pp. 274–289 (2024 May).
https://doi.org/10.17743/jaes.2022.0133.

Sound Terminology in Sonification

TIM ZIEMER*
(tim.ziemer@uni-hamburg.de)

Institute of Systematic Musicology, University of Hamburg, Hamburg, Germany

Sonification research is intrinsically interdisciplinary. Consequently, a proper documentation of and interdisciplinary discourse about a sonification is often hindered by terminology discrepancies between involved disciplines, i.e., the lack of a common sound terminology in sonification research. Without a common ground, a researcher from one discipline may have trouble understanding the implementation and imagining the resulting sound perception of a sonification, if the sonification is described by a researcher from another discipline. To find a common ground, the author consulted literature on interdisciplinary research and discourse, identified problems that occur in sonification, and applied the recommended solutions. As a result, the author recommends considering three aspects of sonification individually, namely 1) Sound Design Concept, 2) Objective, and 3) Evaluation, clarifying which discipline is involved in which aspect and sticking to this discipline's terminology. As two requirements of sonifications are that they are a) reproducible and b) interpretable, the author recommends documenting and discussing every sonification design once using audio engineering terminology and once using psychoacoustic terminology. The appendixes provide comprehensive lists of sound terms from both disciplines, together with relevant literature and a clarification of often misunderstood and misused terms.

0 INTRODUCTION

Sonification is intrinsically interdisciplinary. For example, [1] presents the interdisciplinary circle of sonification, including audio engineering, psychoacoustics, and music, but also cognitive sciences, acoustics/computer music, computer science, linguistics/philosophy, social sciences, product design, data mining/statistics, and many more. They state that physics, acoustics, psychoacoustics, perceptual research, sound engineering, and computer science are "required to comprehend and carry out successful sonification," while respecting that sonification also has facets concerning psychology, musicology, cognitive sciences, linguistics, pedagogies, social sciences, and philosophy. *The Sonification Handbook* [2] contains chapters dedicated to psychoacoustics [3], perception and cognition [4], human-computer interaction [5], and design and aesthetics [6]. David Worrall's *Sonification Design* [7] covers philosophy, computer science, cognitive sciences, neuroscience, psychology, psychoacoustics, and music theory. Many papers deal with the interdisciplinary nature of sonification research [1, 8–11].

Interdisciplinarity is often considered a strength of sonification research due to the potential to understand the overall picture rather than illuminating single fragments [1, 12]. It is illusive to believe that every researcher who deals with sonification is educated in all these disciplines, nor is this a prerequisite. But clearly, authors and readers of sonification literature tend to be educated in some of these disciplines and lack knowledge and competence in others, which is common in interdisciplinary research [13].

Several hurdles have been recognized for interdisciplinarity to unfold in sonification research, due to

a) No common terminology, especially concerning sound [1, 8, 10], and
b) Blurred assignment between disciplines' methods and the sonification's concept, goal, and/or evaluation [1, 7, 9, 10].

Consequently, authors may focus on different aspects and use different, even mistakable, terms to describe their sonification. This can lead to misconceptions that hinder progress in sonification research. Furthermore, this involves the risk that researchers disregard two essentials of sonification discourse, which are *reproducibility* and *interpretability*: While early definitions refer to sonification as "the use of non-speech audio to convey information" [14], newer

*To whom correspondence should be addressed, email: tim.ziemer@uni-hamburg.de. Last updated: October 20, 2023

definitions add the requirements of the sound being a) *reproducible* and b) *interpretable* [5, p. 274; 15]. Here, the term *reproducible* points at the technical implementation of the sonification. This mostly concerns the signal processing of the parameter mapping or the architecture of a model-based sonification. The term *interpretable* stresses out the fact that it is not sufficient if data is encoded in sound. The signal processing is the means, not the objective.

To fulfill the requirements of sonification, the encoded data must be interpretable by a human listener. Sonification designers therefore must consider the transfer from data over signal parameters to auditory perception [16]. The motivation for parameter mapping should always be the perceptual outcome rather than the signal processing itself. Consequently, sonification designers must have both aspects in mind and communicate them to other researchers. This paper is supposed to help sonification researchers meet these requirements by suggesting a common sound terminology for interdisciplinary sonification research and discourse. “According to many in the community, sonification lacks such a comprehensive common language, although there seems to be a shared understanding that it will become necessary to develop one” [17]. The sound terminology for sonification proposed in this paper is an approach to a common sound terminology in sonification to overcome the two hurdles. Note that it is dedicated to scientific publications. Public outreach, popular science, and communication with experiment participants are out of scope of this paper.

[18] stated that promoting interdisciplinarity in sonification research has been addressed rather hesitantly. Fortunately, studies on interdisciplinary research have already identified general problems and proposed respective solutions. In this paper, the author relates four general problems to sonification-specific problems and then transfers the proposed, general solutions to sonification research. The result is a list of recommendations concerning the use of sound terminology and the clarification of whether a discipline contributes to the sonification concept, goal, and/or evaluation.

The remainder of this paper is structured as follows: SEC. 1 presents the literature on interdisciplinary research that formulates four general problems relating to the issues of a) no common terminology and b) blurred assignment between disciplines’ methods and the sonification’s concept, goal, and/or evaluation. Each general problem is explained and then transferred to sonification research based on prominent examples. Finally, the proposed, general solution is applied to sonification, leading to recommendations for sonification researchers. In SEC. 2, all recommendations are discussed, followed by a conclusion in SEC. 3. APPENDIX A.1 provides literature recommendations for lexical study of sound terminology, includes a comprehensive list of terms and definitions describing the sound of sonifications with the vocabulary of audio engineering and psychoacoustics, and reveals important distinctions between terms that are often confused or misused in the sonification literature.

Table 1. Problems and solutions in interdisciplinary research.

No.	Problem	Solution
1	Conflicting Paradigms	Name concepts, practices, and objectives
2	Strong Interdisciplinarity	Name identity criteria
3	Communication Discrepancies	Carry out lexical study
4	Interdisciplinary Semantic Drifts	Name discipline and stick to its terminology

1 INTERDISCIPLINARITY IN SONIFICATION RESEARCH

Literature on interdisciplinarity in research [19–22] was consulted, and four general problems of interdisciplinary research were identified that relate to the issue of a) no common sound terminology and b) blurred assignment between disciplines’ methods and the sonification’s concept, goal, and/or evaluation. The literature identifies additional problems of interdisciplinary research, e.g., concerning education and institutionalization [13] and the professional identity of interdisciplinary researchers [20]. These are out of scope of this review. The four problems and the proposed solutions of interdisciplinary research are summarized in Table 1.

This section describes each general problem and relates it to sonification, based on prominent examples in the literature. Then, the proposed solutions are applied to sonification research, which results in recommendation on the documentation, description of sonifications, and a dictionary of sound terms that can be found in the appendixes.

1.1 Problem 1: Conflicting Paradigms

[20] describes the issues of conflicting paradigms. Conflicting paradigms refer to different viewpoints and interests. One discipline may be interested in the object under investigation, while the second may be interested in its impact on the individual or society, and a third discipline may be interested in the performance of a specific solution approach. For example, engineers, economists, sociologists, and biologists may evaluate measures against human-made climate change differently, despite the meteorological impact. As an approach to a solution, writers and readers need to detect and interpret involved disciplines and their concepts, practices, and objectives consciously.

1.1.1 Conflicting Paradigms in Sonification

Many sonification studies reveal relevant discipline(s) already in the title, like art [23, 24], music [25], speech [26], design [6, 27], aesthetics [6], perception [4, 7], psychoacoustics [3, 28–32], cognition [4], mathematics [26], philosophy [33], and/or signal processing [32]. This informs readers about discipline(s) involved. For example, [26] describes a sonification based on speech synthesis, using additive synthesis and formant filters leading to the perception of vowels. Unfortunately, many of these studies do not point out whether the discipline’s concepts, objects of interest, or methods and measures are being applied.

[26], for example, refers to speech synthesis but not to an evaluation of the sonification using speech quality assessment, like Absolute Category Rating [34]. This gives rise to conflicting paradigms, as it makes a big difference whether the design or evaluation of the sonification comes from the speech domain.

Similarly, it makes a difference, e.g., whether a musical sonification applies music theory and composition techniques to communicate data to a user, or whether data is used as a foundation for a music composition with the objective of an enjoyable listening experience. Furthermore, some studies' titles [35, 36] include musical terms, even though neither the sonification design concept nor the object of interest is musical. These examples of conflicting paradigms are closely related to the general issue of a blurred assignment between disciplines' methods and the sonification's concept, goal, and/or evaluation [1, 7, 9, 10].

1.1.2 Solution: Name Concepts, Objectives, and Practices

The suggested solution for conflicting paradigms is to name concepts, objectives, and practices. This solution can be applied directly on sonification research. Interdisciplinarity within a single sonification study means that different disciplines may be involved in the three aspects of sonification not necessarily in this particular order:

1. Sound Design Concept
2. Objective
3. Evaluation

The sound design concept typically has one source of inspiration but also two requirements, or identities, which are a) reproducibility and b) interpretability [5, 14, 15]. The source of inspiration can come, e.g., from the field of music [37], psychoacoustics [38], design [39], speech [40], Gestalt psychology [41], or signal processing [32]. For example, a musical sonification design concept can consider multiple data streams as multiple melodic lines; a multimodal sonification may seek for coherence with visual design, like one sound being edgy and another being colorful; and a concept based on Gestalt psychology may aim at communicating characteristics of separate objects as features of segregate auditory streams.

The objective of a sonification study can be, e.g., the conceptualization of a sonification design principle [16], setup of a framework for sonification design [42], or evaluation [43], to evaluate the perception of a sonification [44], the performance [31], or user experience [45] of sonification users, or the sonification's impact on the individual [46], society [47], environment [48], etc. Note that the objective is not independent of the sonification design but rather its driving force [49]. While some of these objectives are irrelevant for some disciplines, sonification researchers should always be interested in all these objects of interest, for the bigger picture of sonification. Depending on the objective, different evaluation methods may be appropriate. Evaluation methods can come from the field of psychoacoustics [28], human factors [50], UX-Design [51], experimental psychology [44], Gamification [52], Human-Computer Interaction [53], and others [54].

Recommendation: Authors should always refer to all three aspects of sonification, namely 1) Sound Design Concept, 2) Objective, and 3) Evaluation. Authors should always explain which discipline is involved in which aspect. The three aspects and each involved discipline should be mentioned already in the title or abstract of a manuscript.

1.2 Problem 2: Strong Interdisciplinarity

[19] distinguishes between "weak" and "strong interdisciplinarity." In the case of weak interdisciplinarity, the disciplines share a "proximity to the interests of recognition, which are reflected in pre-discursive consent, terminology, etc." As an example, they describe how electrical engineering and mechanical engineering may be necessary to construct a robot, where there are electric circuits on the one hand and mechanics on the other. In this case, each discipline takes care of a distinct aspect of the robot, but they share similar aims and vocabularies.

In the case of strong interdisciplinarity, it is doubtful whether disciplines agree on the object of interest, methods and measures, and terminology. One example is uninformed source separation of music mixes. Computer scientists may use signal correlation between the source and separated tracks to evaluate the separation quality. Psychoacousticians may let participants rate the similarity of the original and separated tracks to evaluate the separation quality. Audio engineers may speak of *polyphonic music* because the audio track contains sounds of multiple instruments that overlap in time and frequency, while musicologists will refer to it as a *music mix*, because polyphony is considered as a certain way of musical thinking (see APPENDIX A.1).

As a solution, the authors state that "disciplines have to organize themselves within interdisciplinary research frameworks" [19, p. ix]. This means clear identity criteria have to be named, e.g., which discipline is responsible for what and whose methods and terminology is applied when [19, SEC. 2.3.3.2].

1.2.1 Strong Interdisciplinarity in Sonification

As stated above, sonification is intrinsically interdisciplinary. But some studies focus on either side, like the viewpoint of audio engineering [11], human factors [55, 56], psychoacoustics [28, 29], design [27], or arts [23, 24, 33]. Readers of such studies can expect to find the terminology from the respective discipline and results that are relevant to the specific discipline. This can be considered a weak interdisciplinarity from the viewpoint of the specific discipline. The same is partly true for sonification studies that focus on solving a specific problem, like [26, 30]. Here, the data and problem come from one discipline, the *domain science* [11], while the sonification to solve the problem comes from another. The contributions of each discipline are clearly separated. This makes the collaboration easier. Nevertheless, it may be difficult for a sonification researcher to read various papers like these, as the terminologies used to describe the sound may vary across studies, i.e., across involved disciplines. This is because sonification research lacks a common terminology, especially concerning sound [1, 8, 10]. Moreover, sonification researchers are often interested in sonification from a multitude of viewpoints and, thus, strong interdisciplinarity. Examples for interdisciplinary discourse on sonification were already provided in the introduction, like [2, 7].

But even these prominent examples of interdisciplinary sonification research sometimes suffer from the strong interdisciplinarity problem. In a short passage of his book, Worrall mixes terminologies from multiple disciplines, writing that "Mappings were to pitch, amplitude and frequency modulation (pulsing and detuning), filter coefficients (brightness) and onset time (attack)" [7, p. 217]. Here, some terms come from the field of audio engineering (amplitude and pitch modulation, filter coefficients, and attack time) while others come from the field of perception (pitch and brightness). As a consequence, it is not completely clear how the sound is manipulated (audio engineering) or how it may sound (psychoacoustics). The technical and perceptual identities of the sonification design are mixed. In another part of the book, he writes about "pitch, loudness, stereo-location" [7, p. 218]. Clearly, pitch and loudness are perceptual terms coming from the field of psychoacoustics. However, the term "stereo-location" is a bit ambiguous, probably referring to both the perceived source location but also to panning technology in a stereo loudspeaker setup (audio engineering). Even though this term appears to be an audio engineering term, it does not reveal how the stereo location was implemented (amplitude-based panning, like the sine law, tangent law, or Chowning's panning law; time-based panning; or fading-based panning [57, SECS. 7.2 and 9.3]). In these passages, the lack of clear identity criteria is evident: Does the description contribute to a) the reproducibility of sonification or b) the interpretability?Authors need to identify and clearly separate these two identities of a single sonification.

1.2.2 Solution: Identity Criteria

To solve the problem of strong interdisciplinarity, the literature suggests that disciplines organize themselves through identity criteria. In the field of sonification, the above-mentioned solution already partially solves this issue: Authors explicitly associate each involved discipline with either of the three aspects of sonification, namely the sound design concept, objective, and/or evaluation method. This assignment is necessary to reveal whether the sound design concept is musical [4, p. 79], such as using musical scales, reverse scoring, step sequencer, voice leading, or counterpoint; whether the objective is musical, i.e., an enjoyable listening experience [25]; and/or whether the sonification evaluation method comes from the field of musicology, like observation combined with mathematical reasoning, comparative methods, or experiments known from the field of Gestalt psychology, psychoacoustics [58]. Additionally, as discussed above, sonification needs to be a) reproducible and b) interpretable, which is why at least a physical and perceptual description of each sonification are needed. In their parameter mapping sonification review [59], Dubus and Bresin already expressed the need for a "structure for the classification of both physical and auditory dimensions."

Recommendation: Authors should always reflect on their sonification's a) technical reproducibility using physical terms and b) interpretability using perceptual terms. These different viewpoints should never be mixed up.

1.3 Problem 3: Communication Discrepancies

[20] discusses the issue of communication discrepancies. What they mean by communication discrepancies is that one discipline uses terms colloquially, which are specifically defined in another discipline. This can lead to the loss of detail or connotation, or even to a total misconception. The authors describe how "the lack of understanding of some terms by collaborators created an unanticipated need to provide definitions" [20, p. 9] in an interdisciplinary workshop. For example, the term *precision* has a clear mathematical definition in the field of information retrieval, but the same term is often used colloquially in other fields to describe exactness, accuracy, closeness, or the like. As a solution, the authors recommend lexical studies to become aware of discipline-specific meanings and connotations of terms.

1.3.1 Communication Discrepancies in Sonification

Communication discrepancies easily arise when talking about sonification. As discussed above, sonification needs to be a) reproducible and b) interpretable, which is why a physical and perceptual description of each sonification is needed. A common complaint is that authors do not distinguish between physical and perceptual terms [4, p. 64; 31, p. 1077; 59, p. 6; 60, p. 20 and Chap. 2]. Prominent examples can be found when reading different chapters of *The Sonification Handbook* [2]: For example, [12] speaks of "...acoustic variables, such as pitch or loudness" [12, p. 18] even though these are clearly auditory (from a perceptual domain) and not acoustic (from the physical domain) variables. It seems that the authors use these terms colloquially and not in the strict sense of acoustics and psychoacoustics. Likewise, in [61, SEC. "Zipper Noise"], White and Louie write about "signal parameter such as loudness."

David Worrall writes, "[T]hat we (at least in English) so frequently substitute the word 'note' for 'tone', and 'music' for 'score"' [7, p. 18]. Some studies only provide sonification descriptions from either side: "...in a sonification of real-time financial data Janata and Childs (2004) used rising and falling pitch to represent the change in price of a stock and loudness to indicate when the stock price was approaching a pre-determined target..." [4, p. 64]. Obviously, this description of the sonification helps the reader imagine how it may sound. And it clarifies the intention of the sonification mapping. But it does not enable the reader to reproduce it.

Another example can be found in Worrall's book: "Earcons are made by transforming a tone's psychophysical parameters—pitch, loudness, duration and timbre—into structured, non-verbal 'message' combinations" [7, SEC. 2.2.1.3]. This citation shows how an author successfully managed to describe earcons on a perceptual level with psychoacoustic terminology, avoiding a mixture with musical or physical terms. Other studies aim at illuminating both the reproduction and interpretation of a sonification: [32] provides one formula describing the complete signal processing of their sonification and then provides equations for all variables, enabling the reader to reproduce the sonification when the three input variables Δx, Δy, and Δz are given. In the section titled "Psychoacoustics," they describe how the signal processing affects perception, using the psychoacoustical terms *pitch height*, *chroma*, *loudness*, *brightness*, *roughness*, *fullness*, *subjective duration*, *tonalness*, and *harmonicity*, and a reference to *auditory scene analysis*. This paper stringently uses psychoacoustic terminology to describe the sonification perception and signal processing formulas to describe the implementation. What is missing is the explanation of the formula in audio engineering terms.

1.3.2 Solution: Lexical Study

According to the literature on interdisciplinary research, the problem of communication discrepancies can be solved through lexical study. This is important, because the same term can have different meanings or connotations in different disciplines. Several sonification studies already aimed at finding a consistent taxonomy, e.g., concerning design [62, 63], listening modes [64], interaction [65], and auditory display [12, 15]. Other sonification studies have already expressed the need for a consistent sound taxonomy, too [1, 17, 27].

As discussed above, the sound design concept of a sonification can originate, e.g., in musical, psychoacoustical, or Gestalt-psychological considerations and techniques. Consequently, it seems nonsensical to establish a standardized sonification terminology, as it would, e.g., force authors to use a non-musical language to express their musical sonification sound design concept. The same is true for the objective and the evaluation method of the sonification study. However, the situation is different concerning reproducibility and interpretability of sonification. The first clearly requires a technical terminology, e.g., from the field of physics/audio engineering [59], while the second requires perceptual terms. Here, psychoacoustic terminology has been suggested [12, 66]. In their review paper about parameter mapping sonification [59], Dubus and Bresin suggest five perceptual high-level categories *Loudness-Related*, *Pitch-Related*, *Spatial*, *Temporal*, and *Timbral* that can serve to describe sonifications. Following this categorization, the present author carried out a comprehensive lexical study of sound descriptors in the field of audio engineering and psychoacoustics. These can be found in the appendixes.

Recommendation: Authors should describe sonifications using audio engineering terminology for the sake of a) reproducibility and using psychoacoustic terminology for the sake of b) interpretability. A comprehensive list of terms, definitions, and further readings can be found in the APPENDIX A.1.

1.4 Problem 4: Interdisciplinary Semantic Drifts

[21] describes how concepts from one discipline are borrowed by another discipline and then acquire new meanings or connotations, called *Interdisciplinary Semantic Drifts* (ISDs). In contrast to the communication discrepancies discussed above, ISDs include terms that are not used colloquially in either discipline. Still, as long as the reader does not know which discipline's definition of a term is applicable, misunderstandings can appear. The example that the authors provide is the term "abortion," which exhibits a very different connotation in the medical field compared to gender studies. In the first, the term describes a medical or surgical procedure of the intentional ending of a pregnancy, while in the second, it goes along with questions of reproductive rights, bodily autonomy, and the impact of policies and social norms on individuals and communities. The easiest solution to this issue is to name the respective discipline whose terminology is used, stick to it as exactly as possible, and mention whenever the terminology from another discipline is used.

1.4.1 ISDs in Sonification

"The vast majority of the tools and techniques used for the computer synthesis of sound have been developed by composers and engineers engaged in the task of making new music" [7, p. 46]. This circumstance is one of the reasons of ISDs in sound terminology. A typical case of a semantic drift can be found in the introductory explanation chapters of Chowning's and Bristow's book *FM Theory & Applications: By Musicians for Musicians* [67]. As the book title implies, the inventor and implementor of FM synthesis and the Yamaha DX7 explain their audio invention to musicians. This leads to referring to *frequency modulation* as *vibrato*, *modulation depth* as *vibrato depth*, and *modulation frequency* as *vibrato rate*. This aim of translating from the audio domain to the musical domain for a better understanding may be one of the origins of semantic drifts.

Note that, originally, *vibrato* is a musical term. Like many musical terms, it is related to the production mechanism of this characteristic sound: vibrating your finger on the string that you are playing or vibrating the telescoping slide of a trombone back and forth. And what this does to the sound is more than a frequency modulation. As the frequency envelope of string instruments is not flat, a vibrato is always a combination of a coherent frequency modulation and incoherent (individual for all partials) amplitude modulation. Generally speaking, the inclusion of electrical engineering and recording studio technology in the creative process of music making initiated reconsideration of traditional musical terms and introduced signal processing terms to the music domain.

Likewise, the term *pitch* from the music domain underwent a semantic drift in the psychoacoustic domain. In the music domain, pitch refers to the vertical dimension of notes and to musical intervals. It often, but not always, refers to the fundamental frequency. In psychoacoustics, pitch refers to a multidimensional perception consisting of a height dimension, chroma dimension, and strength that are affected by periodicity, (fundamental) frequency and frequency distribution, sound pressure level, the place principle, and phase locking of neurons in the auditory nerve [68].

1.4.2 Solution: Name Which Discipline's Terminology It Is

All the above-mentioned solutions underlined the need to name involved disciplines, ideally already in the title or abstract, and assign them to the sound design concept, objective, or methodology. The present author also identified that the requirement of sonifications to be reproducible and interpretable can be covered by the disciplines audio engineering and psychoacoustics and their respective terminology. However, these are just the most essential disciplines involved.

Whenever you use the terminology from other disciplines, such as music(ology), computer science, cognitive sciences, or the like, say so before you use them. And state why you use them. In his book *Sonification Design*, David Worrall describes that the inclusion of algorithmic music and data-driven compositions in sonification is unfortunate, "because it blurs purposeful distinctions" [7, Sec. 1.9]. It is crucial that authors clearly make those distinctions.

Recommendation: Explicitly name disciplines whose terminology you use; whether you use it to describe the sound design concept, ensure reproducibility of sonification, describe interpretation/perception of the sonification, or reveal the objective of your study or the evaluation method that you use.

The proposed solution has been applied in some sonification studies. For example, Neuhoff writes: "...a primary physical characteristic of a tone is its fundamental frequency (usually measured in cycles per second or Hz). The perceptual dimension that corresponds principally to the physical dimension of frequency is 'pitch', or the apparent 'highness' or 'lowness' of a tone" [4, p. 64]. This statement makes the distinction between two disciplines clear, based on their viewpoints and terminologies.

2 DISCUSSION

The suggested solutions from the literature on interdisciplinary research are directly applicable to sonification. Naturally, the vocabulary of a single discipline is limited. Sometimes, authors need to include the terminology from a second discipline, e.g., to ensure a) reproducibility or b) interpretability of a sonification. The terminology from additional disciplines may become necessary, especially when the sonification design concept is inspired by music, speech, or soundscapes. The lexical study provided in the appendixes does not include these disciplines. Additional research is necessary to identify mistakable terms from these disciplines, provide a list of definitions, and expand the list in Appendix A.3.

Theoretically, a sonification does not require signal processing. A famous example is the sonification of a ball's velocity when rolling down a ramp [69, p. 267]: When placing gut frets or light bells equidistantly above the ramp, such that the ball strikes them when passing, you can hear the ball speed up, because the frequency of bell strikes rises, i.e., the inter-onset interval reduces. Here, the audio engineering terminology needs to be expanded by physical terms to assure reproducibility of the setup. Similarly, the psychoacoustic terminology is not capable of expressing all aspects of sound perception. Here, terms from Gestalt psychology (auditory scene analysis) or soundscape studies may be borrowed. However, this should always be mentioned explicitly.

The sound terminology in sonification expands existing taxonomic studies like [12, 15, 62, 65], which aim at standardizing the sonification terminology. Note that terminology is not the only issue that results from the interdisciplinary nature of sonification. Further research is necessary to establish solid recommendations regarding standardized means for the evaluation of sonifications [28, 51, 52, 54, 70–72]. Note that the two problems of sonification research tackled in this study are only the tip of the iceberg. As discussed above, the interdisciplinarity in sonification research goes far beyond sound terminology. However, especially the division of sonification into the three aspects of 1) sound design concept, 2) objective, and 3) evaluation should help authors structure and illuminate their work in such a way that readers from various disciplines should be able to understand it. Agreeing on a standardized terminology and naming involved disciplines in the title or abstract also makes it easier for researchers to find relevant data and literature from databases and repositories [73].

A conclusion of the *Sound Terminology in Sonification* is that the terminology of at last two disciplines is necessary to document and describe a sonification, namely audio engineering and psychoacoustics. The fact that many sonification researchers do not treat the requirements of reproducibility and interpretability (or implementation and intended sound perception) separately causes misunderstandings and misconceptions, and it may also be the reason why no common audio terminology has been suggested before.

Note that even the best textual description of a sonification cannot replace audio or audiovisual examples. These can be rendered and archived easily with modern technology and should be linked as supplementary material whenever possible, as recommended in the Sonification Report [14] and ICAD conference websites and practiced, e.g., by [2–4, 6, 12, 31, 38, 48, 50, 52, 54, 69]. This allows other researchers to understand the sonification not only through imagination but through experience. Similarly, providing mathematical formulas or pseudo source code in a manuscript [11] or linking to actual source codes in a repository helps researchers understand the implementation and reproduce the sonification.

3 CONCLUSION

In this paper, two problems in sonification research related to the use of sound terminology have been identified, namely a) the lack of a common sound terminology and b) the blurred assignment between disciplines' methods and the sonification's concept, goal, and/or evaluation. Through consultation of literature on interdisciplinary research, these two sonification-specific problems could be related to four general problems of interdisciplinary research. Then, the proposed solutions were applied to sonification research, resulting in recommendations concerning sound terminology and the separation of involved disciplines: Authors should not only name all disciplines involved in their study, but also assign them to the aspect 1) Sound Design Concept, 2) Objective, and/or 3) Evaluation.

Regarding the requirement of a sonification to be a) reproducible and b) interpretable, authors should describe both viewpoints of their sonification, ideally using audio engineering and psychoacoustic terminology, respectively. They should consistently stick to the terminology of the respective discipline for the respective aspect. The lexical study provided in the appendixes helps sonification researchers understand the sound vocabulary of audio engineering and psychoacoustics, avoid common misconceptions, and find further reading. The merit of this study is a standardized sound vocabulary, which can be considered the minimum requirement for the documentation of and discourse about sonification. Further efforts are necessary to expand the lexical study by terms from the field of music, speech, and soundscapes.

4 ACKNOWLEDGMENT

I thank the editor Vesa Välimäki and the three anonymous reviewers who provided a lot of food for thought concerning the title, structure, focus, formulations, and formatting of the paper.

5 REFERENCES

[1] T. Hermann, A. Hunt, and J. G. Neuhoff, "Introduction," in T. Hermann, A. Hunt, and J. G. Neuhoff (Eds.), *The Sonification Handbook*, pp. 1–6 (Logos, Berlin, Germany, 2011). http://sonification.de/handbook.

[2] T. Hermann, A. Hunter, and J. G. Neuhoff (Eds.), *The Sonification Handbook* (Logos, Berlin, Germany, 2011). http://sonification.de/handbook/.

[3] S. Carlile, "Psychoacoustics," in T. Hermann, A. Hunt, and J. G. Neuhoff (Eds.), *The Sonification Handbook*, pp. 41–61 (Logos, Berlin, Germany, 2011). https://sonification.de/handbook/.

[4] J. G. Neuhoff, "Perception, Cognition and Action in Auditory Displays," in T. Hermann, A. Hunt, and J. G. Neuhoff (Eds.), *The Sonification Handbook*, pp. 63–85 (Logos, Berlin, Germany, 2011). http://sonification.de/handbook/.

[5] A. Hunt and T. Hermann, "Interactive Sonification," in T. Hermann, A. Hunt, and J. G. Neuhoff (Eds.), *The Sonification Handbook*, pp. 273–298 (Logos, Berlin, Germany, 2011). http://sonification.de/handbook/.

[6] S. Barrass and P. Vickers, "Sonification Design and Aesthetics," in T. Hermann, A. Hunt, and J. G. Neuhoff (Eds.), *The Sonification Handbook*, pp. 145–171 (Logos, Berlin, Germany, 2011). http://sonification.de/handbook.

[7] D. Worrall, *Sonification Design* (Springer, Cham, Switzerland, 2019). https://doi.org/10.1007/978-3-030-01497-1.

[8] A. de Campo, C. Frauenberger, K. Vogt, A. Wallisch, and C. Daye, "Sonification as an Interdisciplinary Working Process," in *Proceedings of the 12th International Conference on Auditory Display (ICAD)*, pp. 28–35 (London, UK) (2006 May). http://hdl.handle.net/1853/50636.

[9] C. Dayé and A. de Campo, "Sounds Sequential: Sonification in the Social Sciences," *Interdiscip. Sci. Rev.*, vol. 31, no. 4, pp. 349–364 (2013 Dec.). https://doi.org/10.1179/030801806X143286.

[10] J. G. Neuhoff, "Is Sonification Doomed to Fail?" in *Proceedings of the 25th International Conference on Auditory Display*, pp. 327–330 (Newcastle upon Tyne, UK) (2019 Jun.). https://doi.org/10.21785/icad2019.069.

[11] K. Vogt and R. Höldrich, "Translating Sonifications," *J. Audio Eng. Soc*, vol. 60, no. 11, pp. 926–935 (2012 Nov.). http://www.aes.org/e-lib/browse.cfm?elib=16636.

[12] B. N. Walker and M. A. Nees, "Theory of Sonification," in T. Hermann, A. Hunt, and J. G. Neuhoff (Eds.), *The Sonification Handbook*, pp. 9–39 (Logos, Berlin, Germany, 2011). http://sonification.de/handbook/.

[13] X. Ming, M. MacLeod, and J. van der Veen, "Construction and Enactment of Interdisciplinarity: A Grounded Theory Case Study in Liberal Arts and Sciences Education," *Learn. Cult. Soc. Interact.*, vol. 40, paper 100716 (2023 May). https://doi.org/10.1016/j.lcsi.2023.100716.

[14] G. Kramer, B. Walker, T. Bonebright, et al., *Sonification Report: Status of the Field and Research Agenda*, NSF Sonification White Paper (1998 Dec.). http://sonify.psych.gatech.edu/∼ben/references/kramer_the_sonification_report_status_of_the_field_and_research_agenda.pdf.

[15] T. Hermann, "Taxonomy and Definitions for Sonification and Auditory Display," in *Proceedings of the 14th International Conference on Auditory Display*, pp. 1–8 (Paris, France) (2008 Jun.). http://hdl.handle.net/1853/49960.

[16] D. Arfib, J. Couturier, K. L., and V. Verfaille, "Strategies of Mapping Between Gesture Data and Synthesis Model Parameters Using Perceptual Spaces," *Organ.Sound*, vol. 7, no. 2, pp. 127–144 (2002 Aug.). https://doi.org/10.1017/S1355771802002054.

[17] A. Supper, "The Search for the 'Killer Application': Drawing the Boundaries Around the Sonification of Scientific Data," in T. Pinch and K. Bijsterveld (Eds.), *The Oxford Handbook of Sound Studies*, pp. 249–270 (Oxford University Press, New York, NY, 2012).

[18] C. Dayé, "Crafting Sonifications: Levels of Interdisciplinary Exchange in Group Discussions and Their Empirical Assessment," in *Proceedings of the 21st International Conference on Auditory Display (ICAD)*, pp. 54–57 (Graz, Austria) (2015 Jun.). http://hdl.handle.net/1853/54104.

[19] C. F. Gethmann, M. Carrier, G. Hanekamp, et al., *Interdisciplinary Research and Trans-Disciplinary Validity Claims* (Springer, Cham, Switzerlad, 2015). https://doi.org/10.1007/978-3-319-11400-2.

[20] K. Daniel, M. McConnell, A. Schuchardt, and M. Peffer, "Challenges Facing Interdisciplinary Researchers: Findings From a Professional Development Workshop," *PLoS One*, vol. 17, no. 4, paper e0267234 (2022 Apr.). https://doi.org/10.1371/journal.pone.0267234.

[21] Z. Wang, S. Peng, J. Chen, A. G. Kapasule, and H. Chen, "Detecting Interdisciplinary Semantic Drift for Knowledge Organization Based on Normal Cloud Model," *J. King Saud Univ. Comput. Inf. Sci.*, vol. 35, no. 6, paper 101569 (2023 Jun.). https://doi.org/10.1016/j.jksuci.2023.101569.

[22] C. A. Cuevas-Garcia, "Understanding Interdisciplinarity in Its Argumentative Context: Thought and Rhetoric in the Perception of Academic Practices," *Interdiscip. Sci. Rev.*, vol. 43, no. 1, pp. 54–73 (2018 Jan.). https://doi.org/10.1080/03080188.2016.1264133.

[23] M. K. Mardakheh and S. Wilson, "A Strata-Based Approach to Discussing Artistic Data Sonification," *Leonardo*, vol. 55, no. 5, pp. 516–520 (2022 Oct.). https://doi.org/10.1162/leon_a_02257.

[24] T. Metcalf, "Graphical Data Sets as Compositional Structure: Sonification of Color Graphs in *RGB* for Clarinet and Piano," *Leonardo*, vol. 54, no. 3, pp. 329–336 (2021 Jun.). https://doi.org/10.1162/leon_a_01964.

[25] C. Scaletti, "Sonification ≠ Music," in R. T. Dean and A. McLean (Eds.), *The Oxford Handbook of Algorithmic Music* (Oxford University Press, New York, NY 2018). https://doi.org/10.1093/oxfordhb/9780190226992.013.9.

[26] T. Hermann, G. Baier, U. Stephani, and H. Ritter, "Kernel Regression Mapping for Vocal EEG Sonification," in *Proceedings of the 14th International Conference on Auditory Display (ICAD)*, pp. 1–7 (Paris, France) (2008 Jun.). http://hdl.handle.net/1853/49939.

[27] S. D. Monache, N. Misdariis, and E. Özcan, "Semantic Models of Sound-Driven Design: Designing With Listening in Mind," *Des. Stud.*, vol. 83, paper 101134 (2022 Nov.). https://doi.org/10.1016/j.destud.2022.101134.

[28] T. Ziemer and H. Schultheis, "PAMPAS: A PsychoAcoustical Method for the Perceptual Analysis of Multidimensional Sonification," *Front. Neurosci.*, vol. 16, paper 930944 (2022 Oct.). https://doi.org/10.3389/fnins.2022.930944.

[29] S. Ferguson, D. Cabrera, K. Beilharz, and H.-J. Song, "Using Psychoacoustical Models for Information Sonification," in *Proceedings of the 12th International Conference on Auditory Display*, pp. 113–120 (London, UK) (2006 Jun.). http://hdl.handle.net/1853/50694.

[30] T. Ziemer and D. Black, "Psychoacoustically Motivated Sonification for Surgeons," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 12, no. 1 (Suppl 1), pp. 265–266 (2017 Jun.). https://doi.org/10.1007/s11548-017-1588-3.

[31] T. Ziemer, H. Schultheis, D. Black, and R. Kikinis, "Psychoacoustical Interactive Sonification for Short Range Navigation," *Acta Acust. united Acust.*, vol. 104, no. 6, pp. 1075–1093 (2018 Nov./Dec.). https://doi.org/10.3813/AAA.919273.

[32] T. Ziemer and H. Schultheis, "Psychoacoustical Signal Processing for Three-Dimensional Sonification," in *Proceedings of the 25th International Conference on Auditory Displays (ICAD)*, pp. 277–284 (Newcastle, UK) (2019 Jun.). https://doi.org/10.21785/icad2019.018.

[33] P. Polotti and G. Lemaitre, "Rhetorical Strategies for Sound Design and Auditory Display: *A Case Study*," *Int. J. Des.*, vol. 7, no. 2, pp. 67–82 (2013 Aug.). http://www.ijdesign.org/index.php/IJDesign/article/view/1201/577.

[34] A. Raake, M. Wältermann, U. Wüstenhagen, and B. Feiten, "How to Talk About Speech and Audio Quality With Speech and Audio People," *J. Audio Eng. Soc.*, vol. 60, no. 3, pp. 147–155 (2012 Mar.). http://www.aes.org/e-lib/browse.cfm?elib=16212.

[35] E. R. Miranda, "The Music of Particle Collisions," *Nature Phys.*, vol. 12, paper 721 (2016 Aug.). https://doi.org/10.1038/nphys3848.

[36] C. Hayward, "Listening to the Earth Sing," in G. Kramer (Ed.), *Auditory Display: Sonification, Audification, and Auditory Interfaces*, pp. 369–404 (Addison-Wesley, Boston, MA, 1994).

[37] G. Leplatre and S. A. Brewster, "An Investigation of Using Music to Provide Navigation Cues," in *Proceedings of the 5th International Conference on Auditory Display (ICAD)*, pp. 1–10 (Glasgow, UK) (1998 Nov.). http://hdl.handle.net/1853/50713.

[38] S. Schwarz and T. Ziemer, "A Psychoacoustic Sound Design for Pulse Oximetry," in *Proceedings of the 25th International Conference on Auditory Display (ICAD)*, pp. 214–221 (Newcastle upon Tyne, UK) (2019 Jun.). https://doi.org/10.21785/icad2019.024.

[39] C. Frauenberger, T. Stockman, and M.-L. Bourguet, "A Survey on Common Practice in Designing Audio in the User Interface," in *Proceedings of the 21st British HCI Group Annual Conference on People and Computers: HCI...But Not as We Know It (BCS-HCI)*, pp. 187–194 (Lancaster, UK) (2007 Sep.) https://doi.org/10.14236/ewic/HCI2007.19.

[40] F. Grond, T. Bovermann, and T. Hermann, "A Supercollider Class for Vowel Synthesis and Its Use for Sonification," in *Proceedings of the 17th International Conference on Auditory Display (ICAD)* (Budapest, Hungary) (2011 Jun.). http://hdl.handle.net/1853/51572.

[41] S. Barrass and V. Best, "Stream-Based Sonification Diagrams," in *Proceedings of the 14th International Conference on Auditory Display (ICAD)*, pp. 1–6 (Paris, France) (2008 Jun.). http://hdl.handle.net/1853/49945.

[42] D. Reinsch and T. Hermann, "sonecules: A Python Sonification Architecture," in *Proceedings of the 28th International Conference on Auditory Display (ICAD)*, pp. 62–69 (2023 Jun.). https://doi.org/10.21785/icad2023.5580.

[43] N. Degara, F. Nagel, and T. Hermann, "Sonex: An Evaluation Exchange Framework for Reproducible Sonification," in *Proceedings of the 19th International Con-*

ference on Auditory Display (ICAD), pp. 167–174 (Łódź, Poland) (2013 Jul.). http://hdl.handle.net/1853/51662.

[44] J. E. Anderson and P. Sanderson, "Sonification Design for Complex Work Domains: Dimensions and Distractors," *J. Experiment. Psychol. Appl.*, vol. 15, no. 3, pp. 183–198 (2009 Mar.). https://doi.org/10.1037/a0016329.

[45] J. P. Bliss and R. D. Spain, "Sonification and Reliability — Implications for Signal Design," in *Proceedings of the International Conference on Auditory Display (ICAD)*, pp. 154–159 (Montréal, Canada) (2007 Jun.). http://hdl.handle.net/1853/50028.

[46] J. Gonzales and W. Yu, "Auditory Display as an Aid for Prosthetic Hand Manipulation: Preliminary Results," in *Proceedings of the 17th International Conference on Auditory Display (ICAD)*, (Budapest, Hungary) (2011 Jun.). http://hdl.handle.net/1853/51576.

[47] F. Kilander and P. Loennqvist, "A Whisper in the Woods — An Ambient Soundscape for Peripheral Awareness of Remote Processes," in *Proceedings of the 8th International Conference on Auditory Display (ICAD)*, pp. 1–5 (Kyoto, Japan) (2002 Jul.). http://hdl.handle.net/1853/51338.

[48] S. Kalonaris, "Reef Elegy: An Auditory Display of Hawaii''s 2019 Coral Bleaching Data," in *Proceedings of the 28th International Conference on Auditory Display (ICAD)*, pp. 70–77 (Norrköping, Sweden) (2023 Jun.). https://doi.org/10.21785/icad2023.5731.

[49] S. Lenzi and P. Ciuccarelli, "Intentionality and Design in the Data Sonification of Social Issues," *Big Data Soc.*, vol. 7, no. 2, pp. 1–8 (2020 Jul.). https://doi.org/10.1177/2053951720944603.

[50] T. Ziemer and H. Schultheis, "Psychoacoustic Auditory Display for Navigation: An Auditory Assistance System for Spatial Orientation Tasks," *J. Multimodal User Interfaces*, vol. 13, no. 3, pp. 205–218 (2019 Sep.). https://doi.org/10.1007/s12193-018-0282-2.

[51] B. J. Tomlinson, B. E. Noah, and B. N. Walker, "BUZZ: An Auditory Interface User Experience Scale," in *Proceedings of the Extended Abstracts of the CHI Conference on Human Factors in Computing Systems (CHI EA)*, paper LBW096 (Montreal, Canada) (2018 Apr.). https://doi.org/10.1145/3170427.3188659.

[52] T. Ziemer and H. Schultheis, "The CURAT Sonification Game: Gamification for Remote Sonification Evaluation," in *Proceedings of the 26th International Conference on Auditory Display (ICAD)*, pp. 233–240 (Online) (2021 Jun.). https://doi.org/10.21785/icad2021.026.

[53] Z. Obrenovic, D. Starcevic, and E. Jovanov, "Experimental Evaluation of Multimodal Human Computer Interface for Tactical Audio Applications," in *Proceedings of the IEEE International Conference on Multimedia and Expo (ICME)*, vol. 2, pp. 29–32 (Lausanne, Switzerland) (2002 Jul.). https://doi.org/10.1109/ICME.2002.1035366.

[54] T. L. Bonebright and J. H. Flowers, "Evaluation of Auditory Display," in T. Hermann, A. Hunt, and J. G. Neuhoff (Eds.), *The Sonification Handbook*, pp. 111–144 (Logos, Berlin, Germany, 2011). http://sonification.de/handbook.

[55] K. Hinckfuss, P. Sanderson, R. G. Loeb, H. G. Liley, and D. Liu, "Novel Pulse Oximetry Sonifications for Neonatal Oxygen Saturation Monitoring: A Laboratory Study," *Hum. Factors*, vol. 58, no. 2, pp. 344–359 (2016 Mar.). https://doi.org/10.1177/0018720815617406.

[56] C. Nadri, S. Ko, C. Diggs, et al., â€œNovel Auditory Displays in Highly Automated Vehicles: Sonification Improves Driver Situation Awareness, Perceived Workload, and Overall Experience," *Proc. Hum. Factors Ergon. Soc. Annu. Meet.*, vol. 65, no. 1, pp. 586–590 (2021 Sep.). https://doi.org/10.1177/1071181321651071.

[57] T. Ziemer, *Psychoacoustic Music Sound Field Synthesis* (Springer, Cham, Switzerland, 2020). https://doi.org/10.1007/978-3-030-23033-3.

[58] A. Schneider, "Systematic Musicology: A Historical Interdisciplinary Perspective," in R. Bader (Ed.), *Springer Handbook of Systematic Musicology*, pp. 1–24 (Springer, Berlin, Germany, 2018). https://doi.org/10.1007/978-3-662-55004-5_1.

[59] G. Dubus and R. Bresin, "A Systematic Review of Mapping Strategies for the Sonification of Physical Quantities," *PLOS One*, vol. 9, no. 4, paper e96018 (2013 Dec.). https://doi.org/10.1371/journal.pone.0082491.

[60] K. Bijsterveld, *Sonic Skills: Listening for Knowledge in Science, Medicine and Engineering (1920s-Present)* (Palgrave Macmillan, London, UK, 2019). https://doi.org/10.1057/978-1-137-59829-5.

[61] G. D. White and G. J. Louie, *The Auditory Dictionary* (University of Washington Press, Seattle, WA, 2005), 3rd ed.

[62] K. V. Nesbitt, "Comparing and Reusing Visualisation and Sonification Designs Using the MS-Taxonomy," in *Proceedings of the 10th International Conference on Auditory Display (ICAD)*, (Sydney, Australia) (2004 Jun.). http://hdl.handle.net/1853/50848.

[63] K. Enge, A. Rind, M. Iber, R. Höldrich, and W. Aigner, "Towards a Unified Terminology for Sonification and Visualization," *Personal and Ubiquitous Computing*, vol. 27, no. 5, pp. 1949–1963 (2023 Oct.). https://doi.org/10.1007/s00779-023-01720-5.

[64] P. Vickers, "Ways of Listening and Modes of Being: Electroacoustic Auditory Display," *J. Sonic Stud.*, vol. 2 (2012).

[65] V. Goudarzi, "Exploring a Taxonomy of Interaction in Interactive Sonification Systems," in T. Ahram, R. Taiar, K. Langlois, and A. Choplin (Eds.), *Human Interaction, Emerging Technologies and Future Applications III*, pp. 140–145 (Springer, Cham, Switzerland, 2020). https://doi.org/10.1007/978-3-030-55307-4_22.

[66] B. N. Walker and G. Kramer, "Ecological Psychoacoustics and Auditory Displays: Hearing, Grouping, and Meaning Making," in J. G. Neuhoff (Ed.), *Ecological Psychoacoustics*, pp. 149–174 (Elsevier, Amsterdam, The Netherlands, 2004).

[67] J. Chowning and D. Bristow, *FM Theory & Applications: By Musicians for Musicians* (Yamaha Music Foundation, Tokyo, Japan, 1986).

[68] A. Schneider, "Pitch and Pitch Perception," in R. Bader (Ed.), *Springer Handbook of Systematic Musicology*, pp. 605–685 (Springer, Berlin, Germany, 2018). https://doi.org/10.1007/978-3-662-55004-5_31.

[69] T. Bovermann, J. Rohrhuber, and A. de Campo, "Laboratory Methods for Experimental Sonification," in T. Hermann, A. Hunt, and J. G. Neuhoff (Eds.), *The Sonification Handbook*, pp. 237–272 (Logos, Berlin, Germany, 2011). http://sonification.de/handbook/.

[70] B. N. Walker, "Magnitude Estimation of Conceptual Data Dimensions for Use in Sonification," *J. Experiment. Psychol. Appl.*, vol. 8, no. 4, pp. 211–221 (2002 Dec.). https://doi.org/10.1037/1076-898X.8.4.211.

[71] B. N. Walker, "Consistency of Magnitude Estimations with Conceptual Data Dimensions Used for Sonification," *Appl. Cognit. Psychol.*, vol. 21, pp. 579–599 (2007 Jul.). https://doi.org/10.1002/acp.1291.

[72] N. Degara, T. Kuppanda, and F. Nagel, "The Walking Game: A Framework for Evaluating Sonification Methods in Blind Navigation," in *Proceedings of the 4th Interactive Sonification Workshop (ISon)*, pp. 52–57 (Erlangen, Germany) (2013 Dec.).

[73] N. Karam, C. Müller-Birn, M. Gleisberg, et al., "A Terminology Service Supporting Semantic Annotation, Integration, Discovery and Analysis of Interdisciplinary Research Data," *Datenbank Spektrum*, vol. 16, pp. 195–205 (2016 Nov.). https://doi.org/10.1007/s13222-016-0231-8.

[74] W. M. Hartmann, *Principles of Musical Acoustics*, Undergraduate Lecture Notes in Physics (ULNP) (Springer, New York, NY, 2013). https://doi.org/10.1007/978-1-4614-6786-1.

[75] J. G. Roederer, *The Physics and Psychophysics of Music: An Introduction* (Springer, New York, NY, 2009), 4th ed.

[76] J. H. Ginsberg, *Acoustics-A Textbook for Engineers and Physicists: Volume I: Fundamentals* (Springer, Cham, Switzerland, 2018). https://doi.org/10.1007/978-3-319-56844-7.

[77] J. H. Ginsberg, *Acoustics-A Textbook for Engineers and Physicists: Volume I: Applications* (Springer, Cham, Switzerland, 2021), corrected ed. https://doi.org/10.1007/978-3-319-56847-8.

[78] Acoustical Society of America and Institute of Radio Engineers, *Acoustical Terminology* (American Standard Association, New York, NY, 1951).

[79] J. Lefebvre, "Physical Basis of Acoustics," in P. Filippi, D. Habault, J.-P. Lefebvre, and A. Bergassoli (Eds.), *Acoustics: Basic Physics, Theory and Methods* pp. 1–38 (Academic Press, London, UK, 1999).

[80] T. Ziemer, "Source Width in Music Production. Methods in Stereo, Ambisonics, and Wave Field Synthesis," in A. Schneider (Ed.), *Studies in Musical Acoustics and Psychoacoustics*, pp. 299–340 (Springer, Cham, Switzerland, 2017). https://doi.org/10.1007/978-3-319-47292-8_10.

[81] R. Mores, "Music Studio Technology," in R. Bader (Ed.), *Springer Handbook of Systematic Musicology*, pp. 221–258 (Springer, Berlin, Germany, 2018). https://doi.org/10.1007/978-3-662-55004-5_12.

[82] A. Drymonitis, *Digital Electronics for Musicians* (Apress, Berkeley, CA, 2015). https://doi.org/10.1007/978-1-4842-1583-8.

[83] E. J. Heller, *An Experiential Approach to Sound, Music, and Psychoacoustics* (Princeton University Press, Princeton, NJ, 2013).

[84] M. Bruneau, *Fundamentals of Acoustics* (ISTE, London, UK, 2006).

[85] M. Gallagher, *The Music Tech Dictionary: A Glossary of Audio-Related Terms and Technologies* (Course Technology PTR, Boston, MA, 2009).

[86] P. Deymier and K. Runge, *Sound Topology, Duality, Coherence and Wave-Mixing* (Springer, Cham, Switzerland, 2017). https://doi.org/10.1007/978-3-319-62380-1.

[87] E. Sengpiel, "What is Amplitude?" in *Forum für Mikrofonaufnahmetechnik und Tonstudiotechnik* (Sengpielaudio, Berlin, Germany, 2014). http://www.sengpielaudio.com/calculator-amplitude.htm.

[88] A. Schneider, "Fundamentals," in R. Bader (Ed.), *Springer Handbook of Systematic Musicology*, pp. 559–603 (Springer, Berlin, Germany, 2018). https://doi.org/10.1007/978-3-662-55004-5_30.

[89] A. Schneider, "Perception of Timbre and Sound Color," in R. Bader (Ed.), *Springer Handbook of Systematic Musicology*, pp. 687–726 (Springer, Berlin, Germany, 2018). https://doi.org/10.1007/978-3-662-55004-5_32.

[90] B. Truax (Ed.), *Handbook for Acoustic Ecology* (Cambridge Street Publishing, Burnaby, BC, 1999), 2nd ed. https://www.sfu.ca/~truax/handbook2.html.

[91] J. London, *Hearing in Time: Psychological Aspects of Musical Meter* (Oxford University Press, New York, NY, 2004).

[92] P. Guillaume, *Music and Acoustics* (ISTE, London, UK, 2006).

[93] H. Fastl and E. Zwicker, *Psychoacoustics: Facts and Models* (Springer, Berlin, Germany, 2007), 3rd ed. https://doi.org/10.1007/978-3-540-68888-4.

[94] S. A. Gelfand, *Hearing: An Introduction to Psychological and Physiological Acoustics* (Informa, New York, NY, 1990), 2nd ed.

[95] J. Blauert, *Spatial Hearing: The Psychophysics of Human Sound Source Localization* (MIT Press, Cambridge, MA, 1997), revised ed.

[96] W. Aures, "Berechnungsverfahren für den Sensorischen Wohlklang Beliebiger Schallsignale (A Model for Calculating the Sensory Euphony of Various Sounds)," *Acustica*, vol. 59, no. 2, pp. 130–141 (1985 Dec.).

[97] K. Tsuji and S. C. Müller, *Physics and Music: Essential Connections and Illuminating Excursions* (Springer, Cham, Switzerland, 2021). https://doi.org/10.1007/978-3-030-68676-5.

A.1 AUDIO ENGINEERING TERMS

What follows is a comprehensive list of sound terms from the audio engineering domain subdivided into the six high-level categories *Loudness-Related*, *Pitch-Related*, *Spatial*, *Temporal*, *Timbral*, and *Other*. Note, however, that the provided lists of terms cannot replace concerning yourself with audio engineering and its terminology and exchanging with researchers from this field. Several textbooks and chapters provide a comprehensive overview about acoustics, audio engineering, and signal processing, like [61, 67, 74–83].

1 LOUDNESS-RELATED

Amplitude is often referred to as the strength of a signal regardless of its frequency and is measured in decibels of sound pressure level [76, SEC. 1.1.1; 67, p. 30; 61, SEC. "Amplitude"], but it can also refer to particle elongation/displacement, velocity, or acceleration and to voltage rather than power [61, SEC. "Amplitude"; 84, SEC. 1.1.5.].

Amplitude Modulation [61, SEC. "Amplitude Modulation, abbr. AM"].

Attenuation is the amplitude reduction, often measured in decibels of sound pressure level [61, SEC. "Attenuation"].

Attenuator, aka *Pad* or *Loss Pad*, is a device to reduce the signal amplitude [61, SECS. "Attenuator" and "Pad"; 85, p. 149] .

Beats are periodic temporal envelope fluctuations resulting from superposition of waves with similar frequency, which can be produced, e.g., through additive synthesis or amplitude modulation [78, p. 9].

Complex Amplitude is a complex number to modify signal amplitude and phase [86, SEC. 2.2; 84, SEC. 1.1.5.].

Damping is the addition of friction to a resonant system, reducing the magnitude of the resonance. Resistance is the electrical analogue to damping; not to be confused with dampening, which means adding water [61, SECS. "Damp, Damping" and "Dampen"].

Decibel (dB) is a unitless, logarithmic measure of power, like sound intensity, but it is also used to quantify signal strength in terms of squared sound pressure, voltage, particle elongation, velocity (aka velocity level), acceleration, or the like [61, SEC. "Decibel, or dB"; 78, p. 11; 83, SEC. 2.10].

Dry refers to a signal with little reverberation [61, SEC. "Dry"]. When the signal also contains little ambient noise, it is called *Dead* [61, SEC. "Dead"; 85, p. 48]. *Displacement* or *Deflection/Elongation* over time (not amplitude over time) is plotted in audio files in time domain and is proportional to voltage [87].

Dynamic Range can refer to the ratio of the loudest to the softest part, measured in decibels, but also to the signal-to-noise ratio [61, SEC. " Dynamic Range"].

Fade is the gradual change of volume, like fading in or out [61, SECS. "Fade," "Fade In," and "Fade Out"; 85, p. 71].

Gain is the amount of amplification and can be considered a multiplication of sound pressure by a real value [61, SECS. "Amplifier," "Gain," and "Voltage Gain"].

Inverse Distance Law describes that the sound pressure of a point source decays with a factor $p \propto 1/r$, where r is the distance from the source. The *Inverse Square Law* describes the same circumstance but refers to sound intensity $I \propto 1/r^2$ [88, p. 576; 61, SEC. "Inverse Square Law"; 83, SEC. 2.2; 85, p. 99].

Master is a gain control for all output channels at once [61, SEC. "Master"].

Mute means to silence a track or channel [61, SEC. "Mute"].

Particle Displacement is proportional to voltage in audio signals [84, SEC. 1.1.5.].

Normalization means increasing the elongation/deflection/voltage/sound pressure such that the peak equals the largest possible value [61, SEC. "Normalize"; 85, p. 137].

Particle Velocity is the time derivative of particle displacement [84, SEC. 1.1.5.].

Peak is the largest absolute value of an audio signal over a considered time interval [61, SEC. "Peak"; 85, p. 154].

Resonances are eigenfrequencies of a physical system in which it will vibrate stronger and longer when excited with a signal containing these frequencies [61, SEC. "Resonance"].

Root Mean Square of sound pressure or voltage quantifies the average volume [61, SECS. "RMS" and "Volt, Voltage"].

Sound Energy, aka acoustic energy, in watts is the amount of kinetic and potential energy [77, p. 3640], i.e., "variation of energy produced by the acoustic perturbation" [79, SEC. 1.2.5.].

Power is the rate of doing work, measured in watts, and equals voltage times current [61, SECS. "Power" and "Watt"].

Sound Intensity (I) in watts per square meter (W/m^2) [61, SEC. "Intensity"] is the power per area carried by the wave. It is the product of sound pressure and particle velocity [77, pp. 572f] and proportional to the squared sound pressure [84, SEC. 1.1.5; 79, SEC. 1.2.5; 88, p. 568; 78, p. 12].

Sound Power in watts cannot be measured directly and tends to have a complicated relationship to sound pressure [61, SEC. "Sound Power"].

Sound Pressure (p) is the "magnitude of the pressure variation" [84, p. 18].

Sound Pressure Level is the ratio of two effective (root mean square) sound pressures, expressed in dB [76, SEC. 1.3.1]. Weighted sound pressure levels in dB_{AtoC} attenuate certain frequencies before calculating the sound pressure level [76, SEC. 1.3.2; 61, SECS. "dBA" and "A-Weighting"].

2 PITCH-RELATED

Angular Frequency, $\omega = 2\pi f$, is the circumference of the unit cycle times the frequency.

Cent is the $\sqrt[1200]{2}$ and divides the octave into 1,200 equal frequency ratios [61, SEC. "Cents"; 85, p. 31].

Complex Tone is a sound that contains discrete (not continuous) spectral peaks [89; 78, p. 24].

Decade refers to a frequency ratio of 10:1 and is sometimes used to describe the rolloff of a filter, like 20 dB per decade [61, SEC. "Decade"].

Frequency in hertz (Hz), which replaced the earlier unit of cycles per second (cps), is the number of exact repetitions per second [61, SECS. "Cycles per Second, cps," "Frequency," and " Hertz, Hz"; 85, pp. 78, 87].

Frequency Modulation means processing frequencies so that they vary continuously over time [78, p. 10; 85, p. 78].

Fundamental Frequency is the lowest frequency from a complex tone [76, p. 43; 61, SEC. " Fundamental"; 78, p. 8; 85, p. 80].

Period is the reciprocal of frequency and can refer to the duration until a function repeats itself [76, SEC. 1.1.1; 78, p. 7].

Periodicity is the frequency of exact signal repetition, irrespective of the fundamental frequency [76, p. 43].

Pitch Bend is a controller and/or MIDI message to alter the speed of a sound playback, simulating the musical playing technique of pitch bending, known., e.g., from guitars, violins, and trombones [85, p. 158].

Pure Tone, aka *Simple Tone*, has a sinusoidal waveform and contains only one frequency [61, SEC. 'Pure Tone"; 78, p. 24].

Sweep, aka *Chirp*, is a frequency that continuously rises (or falls) as a mostly exponential function of time [61, p. 382].

Wow is a slow (below 5 Hz) frequency modulation caused by speed variation in tape recorders [61, SEC. "Flutter"].

3 SPATIAL

Amplitude-Based Panning means gain weighting between several channels, e.g., according to Chowning's/sine/tangent panning law, with Vector Base Amplitude Panning (VBAP) or Multiple Direction Amplitude Panning (MDAP), or through microphone techniques like Y-X, ORTF, or Blumlein, aka intensity stereo [61, SECS. " Intensity Stereo" and "X-Y Stereo"; 80]).

Depth refers to aims at manipulating the apparent distance of a sound source in stereo and thus, indirectly, its size [61, SEC. "Depth"].

Dichotic, **Diotic**, and **Monotic** presentation refers to headphone signal being individual for each ear, equal for each ear, and presented to one ear only [61, SECS. "Dichotic," "Diotic," and "Monotic"].

Sonic Environment refers to sounds at a certain place and time [90].

Phantom Source is a desired location and extent of an auditory event in stereo setups [85, p. 155].

4 TEMPORAL

Burst is a test signal that typically lasts some milliseconds, like a tone burst and noise burst [61, SEC. "Burst"].

Cycle includes all elements that repeat periodically [78, p. 7].

Echo is a distinct, damped repetition of a signal with a delay of 50 ms or more [61, SEC. "Echo"; 78, p. 10].

Impulse is a short, broadband spike [85, p. 96].

Initial Phase (ϕ_0) is the starting point of an oscillation function $p(t) = \hat{A} \sin(\omega t + \phi_0)$ [31].

Inter-Onset-Interval (IOI) is the duration between two successive note, noise, impulse, or sound onsets [91, p. 4].

Period is the duration of a full cycle of a waveform [85, p. 154].

Phase can describe the argument of an oscillation function or the angle of a single frequency's complex amplitude [86, SEC. 2.2; 57, SEC. 5.1.4].

Transient is a nonrepeating waveform, including note onsets, offsets, decays, and modulation [61, SEC. "Transient"].

5 TIMBRAL

Bandwidth is the frequency region in which a spectrum contains significant energy [76, SEC. 1.5.1; 61, SEC. "Bandwidth"].

Bass is often considered the low-frequency portion of the audible frequency range up to about 200 Hz [61, SEC. "Bass"].

Broadband or *Wideband* refers to a wide distribution of spectral energy [61, SECS. "Broadband" and "Wideband"].

Clipping means that the waveform looks like it was clipped by a pair of scissors, producing harmonic distortion and adding to the perception of roughness, but may also make speech easier to understand in noisy environments [61, SEC. "Clipping"].

Corner Frequency is the frequency of a filter where the amplitude is reduced by 3 dB, like the *Cutoff Frequency* of a high-pass or low-pass filter [61, SECS. "Corner Frequency" and "Cutoff Frequency"; 85, pp. 41, 44].

Duty Cycle is the pulse width per period in pulse width modulation synthesis, affecting the weighting of harmonics [82, p. 16; 85, p. 59].

Extreme Highs are frequencies above 10 kHz.

Extreme Lows are frequencies below 40 Hz [61, SEC. "Distortion"].

Filters include *High*, *Low*, *Stopband/Band Reject/Band Elimination*, and *Bandpass Filters* that let frequencies above, below, outside, and, respectively, between the corner frequencies pass while attenuating the others. *Notch Filters* attenuate; *Resonance/Peak Filters* amplify mostly a narrow frequency region.

Comb Filters exhibit a series of deep notches. *Shelf Filters* have a flat response over one or two large frequency ranges and a gradual slope within the transition range [61, SECS. "Bandpass Filter," "Band Reject Filter," "Brickwall Filter and Low-Pass Filter," "Comb Filter," "High-Pass Filter," "Notch Filter," and "Shelving"; 78, p. 18; 81; 85, pp. 16, 35, 74, 88, 113f, 138f, 154].

Frequency Bands, like octave bands and 1/3-octave bands, group the frequency spectrum into smaller portions [76, SEC. 1.3.3; 61, SEC. "One-Third Octave Filter"].

Harmonic Series are frequencies that are integer multiples of a fundamental frequency [85, p. 86].

Highs are frequencies between 4,000 and 10,000 Hz [61, SEC. "Distortion"; 85, p. 88].

Lows are frequencies roughly between 40 and 300 Hz [61, SEC. "Distortion"].

Mid-Bass means frequencies around 200–300 Hz [61, SEC. "Bass"].

Mid-Range are frequencies between 300 and 4,000 Hz [61, SEC. "Midrange"].

Modulation Index is a ratio that expresses the amplitude ratio of carrier and modulation signals in AM synthesis and the frequency deviation between carrier and the highest or lowest sideband in FM synthesis [61, SEC. "Modulation Index"].

Narrowband means that most spectral energy is concentrated in a narrow frequency band, e.g. 400 cent [61, SEC. "Narrowband"].

Noise can mean an unwanted sound not related to the desired signal or a stochastic, aperiodic process, like *Gray, White, Brown/Red, Pink, Blue, Purple Noise* [61, SECS. "Noise," "One-over-f (1/f) Noise," and "White Noise"; 76, SECS. 1.5.1f].

Overdrive means that the gain drives the signal above the linear operating level of equipment, causing overload distortion [61, SECS. "Operating Level," "Overdrive," and "Overload"].

Passband is the frequency region not rejected by a filter [61, SEC. "Passband"; 85, pp. 151f].

Pulse Width Modulation (PWM) is a technique to produce complex tones [82, Chap. 2].

Q, *Q-Factor* or *Quality-Factor*, is the steepness of a filter in the frequency domain [61, SEC. "Q"; 81].

Slope (aka skirt) is the steepness of a filter in the frequency domain, expressed in decibels per decade or octave [61, SEC. "Slope"].

Sound Synthesis includes techniques such as *Additive/Fourier Synthesis*, *Frequency Modulation (FM) Synthesis*, *Granular Synthesis*, *Ring Modulation/Amplitude Modulation Synthesis*, *Physical Modeling*, *Subtractive Synthesis*, and combinations [61, SECS. "Additive Synthesis," "FM Synthesis," "Ring Modulator," and "Subtractive Synthesis"; 92, SEC. 6.1; 67; 85, pp. 4, 7, 76f, 82, 156f; 74, Chap. 9].

Spectral Centroid is the center of gravity of a magnitude spectrum [80, SEC. 4.3.1].

Spectrum, or *Frequency Spectrum*, is the frequency representation of a time signal, indicating amplitude and phase per frequency over a time frame [61, SEC. "Spectrum"; 78, p. 8].

Synthesizer is a hardware and/or software that electronically produces sound [61, SEC. "Synthesizer"].

(Temporal) Envelope is the amplitude (not displacement/deflection/elongation/voltage) over time. It can also refer to an amplitude or gain function $A(t)$, like *Attack Decay Sustain Release (ADSR)* curves of synthesizers [89; 61, SEC. "Additive Synthesis"; 85, pp. 4, 66f; 90; 88, p. 595].

Treble are frequencies above about 2 kHz [61, SEC. "Treble"].

Waveform is the shape of an oscillation over time, like *Sine*, *Triangle*, *Sawtooth*, *Square*, *Pulse Rain*, *Complex Waves*, and *White* (random) *Noise* [61, SECS. "Complex Wave," "Sine Wave," "Square Wave," and "Waveform"].

6 OTHER

Audio Effects is signal processing to modify an audio signal, like *Chorus*, *Compressor*, *Delay*, *Distortion*, *Equalizer*, *Filter*, *Flanger*, *Phaser*, and *Reverb Wah Wah* [61, SECS. "Chorus," "Effects," "Equalizer," "Flanging," " FX"; 92, SEC. 6.3.1. and p. 169; 85, pp. 32, 37, 41, 50, 67, 156].

Carrier is a signal about to be modulated by another signal, the *modulator*, e.g., in AM and FM synthesis [85, p. 29].

Low-Frequency Oscillator (LFO) has a (fundamental) frequency way below 20 Hz and is used to modulate audio material [85, p. 110].

Musical Instrument Digital Interface (MIDI) is a protocol to transmit control data between audio (and lighting) equipment [85, pp. 122ff].

Open Sound Control (OSC) is a network protocol for communication between multimedia devices [85, p. 145; 82, Chap. 8].

A.2 PSYCHOACOUSTIC TERMS

What follows is a comprehensive list of sound terms from the perception/psychoacoustics domain subdivided into the six high-level categories *Loudness-Related*, *Pitch-Related*, *Spatial*, *Temporal*, *Timbral*, and *Other*. Note, however, that the provided lists of terms cannot replace concerning yourself with psychoacoustics and its terminology and exchanging with researchers from this field. Several textbooks and chapters provide a comprehensive insight into psychoacoustic terminology, methods, and paradigms, like [57, 68, 75, 80, 83, 88, 89, 93–95].

1 LOUDNESS-RELATED

Fluctuation Strength refers to the intensity of loudness fluctuation caused, e.g., by amplitude or frequency modulation or beats. The unit of fluctuation strength is vacil, but it is also often given in percent [93, Chap. 10; 61, SEC. "Beats"].

Loudness is the perception that lets listeners order sounds from soft to loud, i.e., the subjective strength of an audio signal that depends on amplitude and frequency but even on spatial and temporal attributes [78, p. 22; 76, SEC. 1.3.2; 61, SECS. "Loudness" and "Sone"].

Masking means that the presence of one sound inhibits the ability to hear another simultaneous (*Simultaneous Masking*), shortly preceding (*Pre-Masking*), or subsequent (*Post-Masking*) sound. Typically, loud sounds mask soft sounds; low frequencies mask higher frequencies [61, SEC. "Masking"; 57, SEC. 4.3].

Phon is the unit of loudness level and is a frequency-compensated decibel scale (aka *Fletcher-Munson-Curve* or *Equal-Ludness-Contour*) but not a unit of loudness [74, p. 127; 78, p. 23; 85, pp. 76, 156].

Sone is a measure of loudness, where a 1-kHz pure tone at 40 dB equals 1 Sone, while a sound twice as loud equals 2 Sone, etc. [78, p. 22].

Specific Loudness is the loudness per critical frequency band. Their integral is the overall loudness [93, Chap. 8].

Threshold of Hearing, aka *Absolute Threshold* or *Minimum Audible Field* (MAF), is the sound pressure level at the eardrum required to be just audible, which is highly frequency dependent [61, SECS. "Minimum Audible Field, MAF" and "Threshold of Hearing"].

2 PITCH-RELATED

Absolute Pitch, aka *Perfect Pitch*, refers to a listener's ability to assign sounds their corresponding position or note on a musical scale [61, SEC. "Absolute Pitch"].

Combination Tones, aka *Phantom Tones/Tartini Tones*, are audible frequencies produced through nonlinearities in the human hearing system that are not physically present in the sound, like *Combination Tones* $f_1 + f_2$, *Difference Tones* $f_2 - f_1$, and *Cubic Difference Tones* $2f_1 - f_2$ [94, Chap. 12; 85, pp. 52, 206f; 68, SECT. 31.6.3].

(In)Harmonicity refers to whether simultaneous frequencies seem to exhibit a harmonic series or not, which can affect simultaneous stream segregation and, therefore, pitch, consonance, timbre, and tonal fusion [68; 57, SEC. 4.5.2].

Mel is the unit of pitch, where 1,000 Mel = pitch evoked by a 1-kHz pure tone at 40 phon (or dB) [61, SEC. "Mel"; 78, p. 22].

Missing Fundamental, aka *Residual Pitch* or *Virtual Pitch*, means that the perceived pitch of a sound equals the pitch of a pure tone whose frequency is not present in the sound [93, Chap. 5; 68; 61, SEC. "Residue"].

Pitch is the perception that lets listeners order sounds from low to high (height) and that includes octave identity/equivalence (chroma) [68; 93, Chap. 5; 78, p. 22; 94, Chap. 12].

Pitch Strength means how distinct a sound's pitch is. It is, e.g., low for clicks, higher for inharmonic series, and even higher for pure tones in the midrange frequency and with sufficient loudness [93, Chap. 5; 68].

Shepard Illusion is pitch perception with clear chroma but ambiguous height and can be created through Shepard tones [88; 68; 57, SEC. 4.5.4.1].

3 SPATIAL

Ambience refers to acoustic qualities of a listening space, including echoes, reverberation, and background noise, and may provide a certain auditory "atmosphere" [61, SEC. "Ambience"].

Apparent Source Width, aka *Perceived Source Extent*, refers to the perceived size of a sound source, sometimes categorical, sometimes indicated in degrees [80; 57, Chap. 6.2.2].

Auditory Event Angle/Distance/Location refers to the perceived source angle/distance/location irrespective of the physical position [80; 57, SEC. 4.4; 95].

Precedence Effect, aka *Haas Effect* or *Law of the First Wavefront*, means that the first wavefront arriving at the ears determines the perceived source location, even in the presence of louder reflections and reverberation [61, SEC. "Haas Effect"; 57, SEC. 4.4.5; 85, p. 85].

Soundscape refers to how the sonic environment is perceived and understood [90].

4 TEMPORAL

Rhythm is a temporal pattern of one auditory stream and somewhat related to fluctuation strength and subjective duration [93, Chap. 13].

Subjective Duration seems quite proportional to physical duration above 100 ms. But temporal masking may affect subjective duration and, e.g., 50-ms sound longer than one half of 100 ms [93, Chap. 12].

Temporal Masking contains pre-masking of up to 30 ms, the *Overshoot* phenomenon at masker onsets, simultaneous masking, a short sustain of 5 ms, and post-masking up to 100 ms [93, SEC. 4.3.1.5].

5 TIMBRAL

Brightness (from warm/dull to bright/shrill) is considered the most prominent feature in timbre evaluation and seems closely related to the spectral centroid of a signal [57, p. 15; 88, p. 595].

Clangorous is a continuous loud noise of something being hit or rung, often metallic [68, p. 704].

Fullness, aka *Richness*, sometimes *Volume*, is mostly affected by the spectral distribution and bandwidth and can be increased, e.g., by adding or modulating frequencies [89, 32].

Noisiness indicated in Noy refers to perceived noise loudness, not to the degree of tonalness [61, SEC. "Noy"].

Percussive to *Mellow* is the degree of how impulsive the excitation of a source sounds, where hammered/struck often sound more percussive than plucked, and even more than blown/stringed [89, p. 701; 88, p. 595].

Roughness in asper is perceived when two or more than 16 but less than one critical bandwidth apart [61, SEC. "Beats"; 96; 89].

Sharpness is equal or related to *Brightness* but has the measure acum that can be approximated by psychoacoustic models [96; 93, Chap. 9].

Sound Color, aka *Tone Color*, refers to the timbre of quasi-stationary sounds, excluding attack transients and decay [89; 61, p. 393].

Timbre includes spectral and short-term temporal aspects of sound other than pitch and loudness [89].

Tonality, aka *Tonalness*, ranges from tonal to noisy and is affected by frequency, harmonicity, periodicity, and bandwidth and considered a parameter of sensory euphony [57, p. 41; 89, p. 700 and SEC. 32.2.3; 96].

6 OTHER

Auditory refers to the psychological representation of sound [57, Chap. 4; 88, p. 586].

Auditory Stream is the auditory counterpart of a visual object, i.e., aspects of sound that seem to belong together [57, SEC. 4.5; 89, SEC. 32.2.5].

Critical Bandwidth is a region of 1.3 mm on the basilar membrane in which neurons collectively process sound [61, SEC. "Critical Band"; 88, p. 584].

Just Noticeable Difference, aka *Difference Limen*, is the increment in a stimulus that is just noticed in a specified fraction of trials [61, SEC. "Just Noticeable Difference, JND"; 78, p. 23].

Perception is a psychological impression often resulting from sensation plus cognitive processing in the brain [88, SEC. 30.1.2]

Sensation can be considered an inter-subjective primal perception before the subjective interference through context, experience, preference, etc. Psychoacoustic models predict sensation and not perception [57, Chap. 4; 93, SEC. 1.4].

Sensory Euphony, aka *Sensory Pleasantness* or *Psychoacoustical Annoyance*, tends to be considered as a nonlinear combination of *Loudness*, *Roughness*, *Sharpness*, and *Tonality* [93, Chap. 9].

Sound Object is the smallest self-contained element of a soundscape [90].

A.3 DANGER OF CONFUSION

Terms with different meanings in different disciplines are identified. The following list shall help researchers avoid mistakable expressions:

A-B Stereo ≠ A-B testing ≠ ABX Test. The first refers to a spaced microphone recording technique, the second to the direct comparison of two versions of the same signal, and the latter is an experiment in which participants guess whether a test stimulus X was equal to reference stimulus A or B [80; 61, SECS. "A-B Stereo," "A-B Testing," and "ABX Test"].

Acoustic ≠ Auditory. The first refers to physics, and the second to perception of sound [95].

Acoustic Recording ≠ Sound Recording. The first refers to recording without the use of electricity, and the second to all sound recording [61, SEC. "Acoustic Recording"].

Amplitude Modulation ≠ Tremolo. The first refers to analogue or digital signal processing, and the latter to an articulation technique, e.g., of string instruments [85, pp. 7, 8].

Audio Track ≠ Audio Channel. Each audio channel can contain a mixture of signals from one or more audio tracks, which can be recorded and manipulated separately, before routing them to discrete output channels. [61, SECS. "Multitrack" and "Tracking"; 85, pp. 31, 132].

Auditory Stream ≠ Audio Stream. An auditory stream is comparable to a visual object, i.e., it includes sound aspects that are perceived as being part of the same. An audio stream is the broadcast of one or multiple audio channels. [61, SEC. "Streaming"].

Bandwidth in terms of the audio signal ≠ **Bandwidth/Throughput/Data Rate** of a network connection or interface [85, p. 16].

Bar in physics ≠ **Bar** in music. In physics, Bar is a unit of pressure, and a bar is a uniform elastic rod or beam with irregular eigenfrequencies. The latter is a segment of music often having a uniform duration, number of beats, and structure of accentuation, probably named after the vertical bar lines in musical scores [97, SECS. 2.3.4 and 3.1.3, pp. 393, 398].

Beats in physics ≠ **Beats** in music. The first refers to envelope fluctuation caused by superposition of similar frequencies, and the latter to the pulse, the rhythm section of a piece of music, or the instrumental, e.g., of hip-hop music [85, p. 19; 97, SECS. 2.3.4 and 3.1.3, pp. 393, 398].

Current ≠ Charge. While the current is a property of a circuit, the charge flows. There is no such thing as a current flow. [61, SECS. "Ampere, abbr. Amp or A" and "Charge"]

dB ≠ $\mathbf{dB_{SPL}}$ ≠ dBA. The first is a power level with arbitrary reference, and the second uses a reference pressure of 20 μPa, just like the third one, which also adds a frequency weighting [61, SECS. "A-Weighting"; 85, pp. 46, 49].

Depth ≠ Auditory Distance ≠ Source Distance. The first refers to stereo methods to affect auditory distance, which is the perceived source distance, while source distance refers to the physical distance of a source [61, SEC. "Depth"; 95].

Delay as an audio effect ≠ **Delay** of a network connection or interface. The first is a single or a number of damped repetitions, and the latter is latency of signal transfer [85, pp. 50, 108].

Elongation/Deflection/Voltage ≠ Amplitude. The first refers to the course of the time signal, and the second to the peak of the absolute time signal [87].

Frequency/Fundamental Frequency ≠ Pitch. The first is physical, and the second is a perceptual quality affected strongly by the first but also by periodicity and sound pressure level [89].

Frequency Modulation ≠ Wow ≠ Vibrato. Frequency modulation describes periodic, continuous frequency variation through signal processing; wow is a frequency modulation resulting from tape speed alterations; and vibrato is a musical instrument playing technique where frequencies and amplitudes of complex tones vary continuously and regularly [61, SEC. "Wow"; 78, pp. 10, 25].

Jitter in speech ≠ **Jitter** in a network connection or interface. The first describes micro-period fluctuations in spoken vowels, and the latter describes timing variations in interfaces or networks [61, SEC. "Jitter"; 85, p. 102].

Loudness ≠ Volume ≠ Sound Pressure Level/Amplitude/Gain. The first is a perceptual quality, the second is a colloquial term loosely related to the others, and the third refers to physical quantities [89; 61, SECS. "Amplitude," "Level," and "Voltage Gain"; 76, SEC. 1.1.1; 78, p. 12; 85, pp. 81, 110].

Meter in music ≠ **Meter** as a measuring device ≠ meter as the SI unit of length [61, SEC. "Meter"].

Monitor in audio (aka *Studio Monitor* or *Stage Monitor*) ≠ **Monitor** in computer science (visual display) [61, SECS. "Monitor" and "Near-Field Monitor"; 85, pp. 127, 133].

Note ≠ Tone. The first is a sign indicating pitch and/or duration of a tone or sound. The second is a sound having a pitch [78, p. 25].

Overshoot in audio ≠ **Overshoot** in perception. The first means that equipment like compressors or limiters responds imperfectly to short transients, while the latter means a brief increase of the masking threshold at sound onsets [61, SEC. "Overshoot"; 57, SEC. 4.3.1.].

Polyphonic in audio engineering $\neq$ **Polyphonic** in music. The first means the overlap of multiple voice, e.g., in synthesizers and music mixes. The second is a horizontal thinking (like counterpoint), i.e., multiple voices playing melodies that result in consonant intervals on pulse, except when the intention is to create tension [75, Chap. 5; 89, pp. 711f].

Overtone $\neq$ **Harmonic** $\neq$ **Partial**. The fundamental frequency is usually the lowest frequency and the first partial, and it may be the first harmonic of a spectrum. The first overtone is the second partial. It is also the second harmonic, if the signal is a complex tone, i.e., if all partials are part of a harmonic series. [61, SECS. "Overtones" and "Partials"; 78, p. 25; 85, pp. 86, 147].

Sharp in music $\neq$ **Sharp** in psychoacoustics $\neq$ **Sharp** in audio engineering. In music, sharp is a rise in pitch, e.g., by one half step, as determined by the accidental "♯." In psychoacoustics, it is a timbral quality, related to brightness. In audio engineering, filters' *Slope/Rolloff/Attenuation/Rejection Rate/Q-Factor* are also called sharpness [61, SEC. "Sharp"; 96; 94, p. 15; 89; 97, SEC. 3.6].

Sound Velocity $\neq$ **Particle Velocity**. The first is a particle's velocity when moving around its equilibrium position, and the second is the propagation speed of a sound wave (perturbation).

Spectral Centroid $\neq$ **Brightness**. The first is the center of gravity of a magnitude spectrum, and the latter an aspect of timbre perception [80, SEC. 4.3.1].

Stereo sometimes $\neq$ **Stereo**. While stereo tracks in digital audio workstations often refer to two channels, one for each loudspeaker in a stereo triangle, stereo can also refer to loudspeaker systems using more than two loudspeakers [61, SEC. "Stereophonic"; 80, Chap. 7].

Tone in physics $\neq$ **Tone** in music. In physics, tone refers to a pure or complex tone, while in musical terms, it can also refer to tone color or even timbre [61, SEC. "Tone"].

Voice in speech $\neq$ **Voice** in music. The first always refers to the human voice, and the second can refer to different instruments or registers [89, pp. 711f].

Timbre $\neq$ **Sound Color/Tone Color**. Timbre includes short-term temporal aspects, like initial transients and decays, while tone color/sound color mostly refers to the spectrum of the quasi-stationary part of sounds [89].

THE AUTHOR

Tim Ziemer

Tim Ziemer is an interim professor for systematic musicology at the University of Hamburg. He gained research experience in musicology at the University of Hamburg, Germany, in computer science at the National Institute of Informatics (NII), Tokyo, Japan, and in medical informatics and cognitive sciences at the University of Bremen, Germany. He is a board member of the International Community for Auditory Display (ICAD), review editor for *Frontiers in Signal Processing*, and co-editor for the *International Journal of the Informatics Society*, Japan. He co-organized several events, like the International Summer School in Systematic Musicology (ISSSM) 2012, Sound and Music Computing (SMC) Summer School 2016, and Interactive Sonification Workshop (ISon) 2022. His research interests include psychoacoustics, recording studio technology, sonification, and hearing mechanisms from mosquitoes to humans.