# Classifying soundscapes using a multifaceted taxonomy

Christopher Trudeau
McGill University, School of Information Studies, Montréal, Canada.

Catherine Guastavino
McGill University, School of Information Studies, Montréal, Canada.

**Summary**

Multiple organizational principles have been proposed to classify information in the field of information studies. Hierarchical classification systems are the most familiar, but they require relational consistency between the different levels and mutual exclusivity across classes. Attempts to organize soundscapes within hierarchical classification are challenging for these two reasons. Listening test results suggest that people cross-categorize everyday sounds according to the sound source, the action generating the sound, and/or the context in which these sounds are perceived. Moreover, soundscapes are complex sound scenes with multiple, concurrent sound sources. This paper proposes a faceted taxonomy that reflects the way people cross-categorize sounds. Five facets are used in the taxonomy, each reflecting a fundamental characteristic of the sound: agent, source, action, context, and acoustic features. Within some facets, a hierarchy of terms can be established, and thus the facets must allow for a hierarchical relationship between the terms. This taxonomy will be implemented using a relational database (MySQL) and a web interface for users to navigate the content and structure of the taxonomy. The proposed database will be populated using existing sound datasets, such as freesound.org. In this taxonomy, a soundscape can be described as a set of sounds, eliminating the challenge of isolating a soundscape to one placement in a taxonomy. We will analyze the free format verbal descriptions accompanying each sound in terms of the facets of the proposed taxonomy, and explore new facets from emerging categories. Thus, the proposed taxonomy supports the ongoing efforts to standardize and report soundscape research by offering a way to systematically describe a given soundscape.

**PACS no. 43.50.Qp, 43.66.Lj**

## 1. Introduction

Categorization is a critical process to make sense of the world around us by dividing it into meaningful categories. Everyday sound categorization allows us to make inferences about the presence of sound sources, agents and actions and subsequently guide action (e.g. avoid an approaching car or attend to a crying baby).

Classification schemes have been proposed to model categorization principles that emerge from empirical and laboratory research. However, current taxonomies have some difficulty representing cross-categorization: our ability to categorize sounds differently according to the situation to infer particular information. For example, the sound of a door slamming, can be categorized based on the source (*e.g. door sound*), or the action (e.g. *someone is mad and slammed the door*) or based on the agent (e.g. *Joe left*).

This paper proposes a faceted taxonomy of sounds to account for cross-categorization. Here, the facets represent the categorization principles that people use, while allowing for a supplementary hierarchical structure. The goal of the proposed taxonomy is to support ongoing research on everyday sounds and soundscapes, as well as to provide a toolkit that can assist urban planners in their consideration of the sonic dimension.

### 1.1. Literature review

Everyday sound categorization has garnered increased research attention within the ecological approach to auditory perception [1] and in the field of soundscape research. In contrast to artificial, synthetic sounds, e*veryday sounds* are defined as sounds occurring in real life environments [2]; they are also referred to as *environmental sounds* in the literature (e.g. [3], [4], [5]) or *domestic sounds* for everyday sounds typically heard inside the home [6]. A variety of everyday sound categorization principles have been reported in the literature (e.g. [5], [7], [6], [8], [9]). There is converging evidence that people categorize everyday sounds based on sound source, the agent or the action producing sounds (see [10] for a review).

As well as cross-categorizing, we also build hierarchies within these categories. Results from free sorting tasks of everyday sound recordings indicated different levels of abstraction in everyday sound categorization, consistent with Rosch's theory of natural categories grouped into superordinate level (e.g. *furniture*), basic level (e.g. *chair*), and subordinate level (e.g. *office chair*).

Research on the categorization of urban soundscapes indicates that complex sound scenes are also subject to cross-classification. A study on soundscape conceptualization found that city users categorized urban sounds according to four main categories: human sounds, traffic sounds, natural sounds and music [11]. Studies using free-sorting tasks of soundscape recordings indicate that categories are based on the presence or absence of human activity [10].

While there is a general convergence in categorization research toward a set of fundamental factors, there is still a need for a standard classification scheme of sounds. Gaver proposed categories of everyday sounds according to the state of the sound source – liquid, gas, solid or mechanical [7] at a superordinate level (e.g. *liquid*), then by actions at a basic level (e.g. *leak*), then by source at a subordinate level (e.g. *faucet*). Guyot et al., 1997, focusing on domestic sounds, found similar evidence that people cross-classify based on the excitation (e.g. *mechanical or electrical*), the action (e.g. *rubbing*, *scratching*) and the source (e.g. *dishes*, *Velcro*). Salamon, Jacoby and Pablo Bello, recently proposed a taxonomy of urban sounds including a combination of sound sources and sound-generating actions [13]. Sounds within an urban acoustic environment are first divided according to four categories: human, nature, mechanical and music. Subordinate to these are high-level semantic classes (e.g. *voice*) and then specific sounds (e.g. *laughter*, an action, or *children*, a source).

Taxonomies for more complex sound scenes have also been proposed. Specifically, Brown, Kang and Gjestland proposed a taxonomy that includes different categorization principles at different levels [13]. They first distinguish by location – indoor or outdoor – and then by type – urban, rural, wilderness or underwater. The third level in the taxonomy is a function of whether or not the sound is created by human activity. Finally, the lowest branches and leaves offer some general and specific sound sources.

### 1.2. Review of taxonomies

A taxonomy is a semantic classification scheme that provides a knowledge map of its domain. In order to meet these criteria, a taxonomy should be comprehensive, predictable and navigable [14].

There are many types of taxonomies available, each with their own strengths and challenges. The classification systems analyzed in the above literature review are all hierarchical taxonomies. This type of taxonomy is generally pre-coordinated and enumerative, that is, the domain is mapped before the items are classified and the mapping is exhaustive of the entire domain. Furthermore, hierarchical taxonomies have the following four characteristics: inclusiveness, relational consistency, inheritance and mutual exclusivity [14]. Inclusiveness implies that parent categories include all of the elements of the subcategories. Relational consistency holds that the relationship between each level is of exactly the same kind. In other words, there is a consistency in the divisions throughout the entire taxonomy. Inheritance holds that the subordinate categories inherit the attributes of the superordinate categories. Finally, mutual exclusivity holds that a single item must be placed unambiguously in one branch of the hierarchy.

In particular, the principle of mutual exclusivity is difficult to enforce for sounds and soundscapes. There is abundant evidence in the literature reviewed that people cross-classify sounds according to several factors. Outside of a laboratory setting, sound scenes are complex and typically involve multiple sound sources concurrently. To some extent, current taxonomies try to account for cross-classification by using sound sources at one level of the taxonomy and sound events at another level, or even a mixture of these at the same level. The result is a taxonomy that presents the user with difficulties during the classification process. Furthermore, the size and potential for unending growth mean that the taxonomy is ill suited to act as a knowledge map. The field of information science offers a taxonomy that is capable of meeting the specific challenges posed by the classification of sounds.

## 2. Faceted taxonomies

Originally proposed by S.R. Ranganathan, faceted classification breaks with the pre-coordinated and enumerative model of hierarchical classification [15]. Facets are the fundamental dimensions (or attributes) of the item that is being classified [14]. In order to avoid confusion, each facet needs to be orthogonal to every other facet in the taxonomy – that is, the attributes of one facet cannot overlap with those of another. For example, a physical item might have colour and texture as two separate facets because there is no conceptual overlap. As the knowledge domain gets smaller, the facets can be more specific and more detailed.

Prior to the development of the digital environment, the opportunities for implementing faceted taxonomies were limited. In particular, libraries need to have a unique location for each resource, even when there are multiple copies of these. However, with the advent of digital technology, faceted taxonomies are able to leverage the use of metadata to aid in implementation [14]. As well, replication costs are nil, making it possible to list the same item in several locations. Therefore, digital technology makes the use of a faceted taxonomy easier.

Through the use of multiple facets, a comprehensive description of a single sound can be created. In this way, the challenge of classifying sounds that are being cross-categorized is minimized because these factors of categorization can be turned into facets.

The drawback of a faceted taxonomy is that it is inherently difficult to visualize. When the number of facets is limited to two (sometimes called a matrix) or even three, it is possible to visualize the content of a facet in 2-space or 3-space. Taxonomies with more than three facets present significant challenges in this area. That said, web technology offers a platform that can be used to visualize the content of a faceted taxonomy.

### 2.1. Implementation

Given the preceding discussion, it is now possible to propose a faceted taxonomy for urban sounds. The subject of this taxonomy will be sound events and not soundscapes for two important reasons. First, combining both sounds and soundscapes in a single taxonomy complicates the process of populating the classifier. Second, any given acoustic environment can be described by its set of sound events. Thus, we can fully describe any given sound scene by listing its component sounds.

The research on sound categorization can be used to map the knowledge domain and inform our choice of facets. Thus, the proposed facets for isolated sound events are: agent, source, action and location. The agent is the human or animal causing the activity that generates the sound (e.g. *child*). The source is the physical material whose vibrations are at the origins of the sound (e.g. *toy*). The action (e.g. *falling*) is the cause of the excitation of the physical material. Finally, the location is a description of the immediate context in which the sound event occurs (e.g. *playroom*).

Each facet will be represented as a single table of a MySQL database, with the arrangement of the items by primary key. The names of the tables mostly correspond to the facet names: *location*, *agent*, *src* (source) and *act* (action). The presentation of the sounds and the taxonomic structure can be done through a web interface, allowing some flexibility to the user to engage and interact with the data.

Implementation into a database requires that we determine the entities that the database needs and what relationship these entities have with each other. Agent, source, action, and location entities can each be attached to multiple sounds. However, each sound can only be attached to one of these. Therefore, we have a one-to-many relationship between the sound event entity and the facet entities. As such, each facet is simply created as a separate entity in the database. Finally, one entity representing the sound event is used to connect each of the facets of the taxonomy together.

The database contains two entities that help to track information about the sounds, but that do not represent facets: *setting* and *origin*. The setting entity is used to distinguish between the acoustic environments that are defined by the Brown, Kang

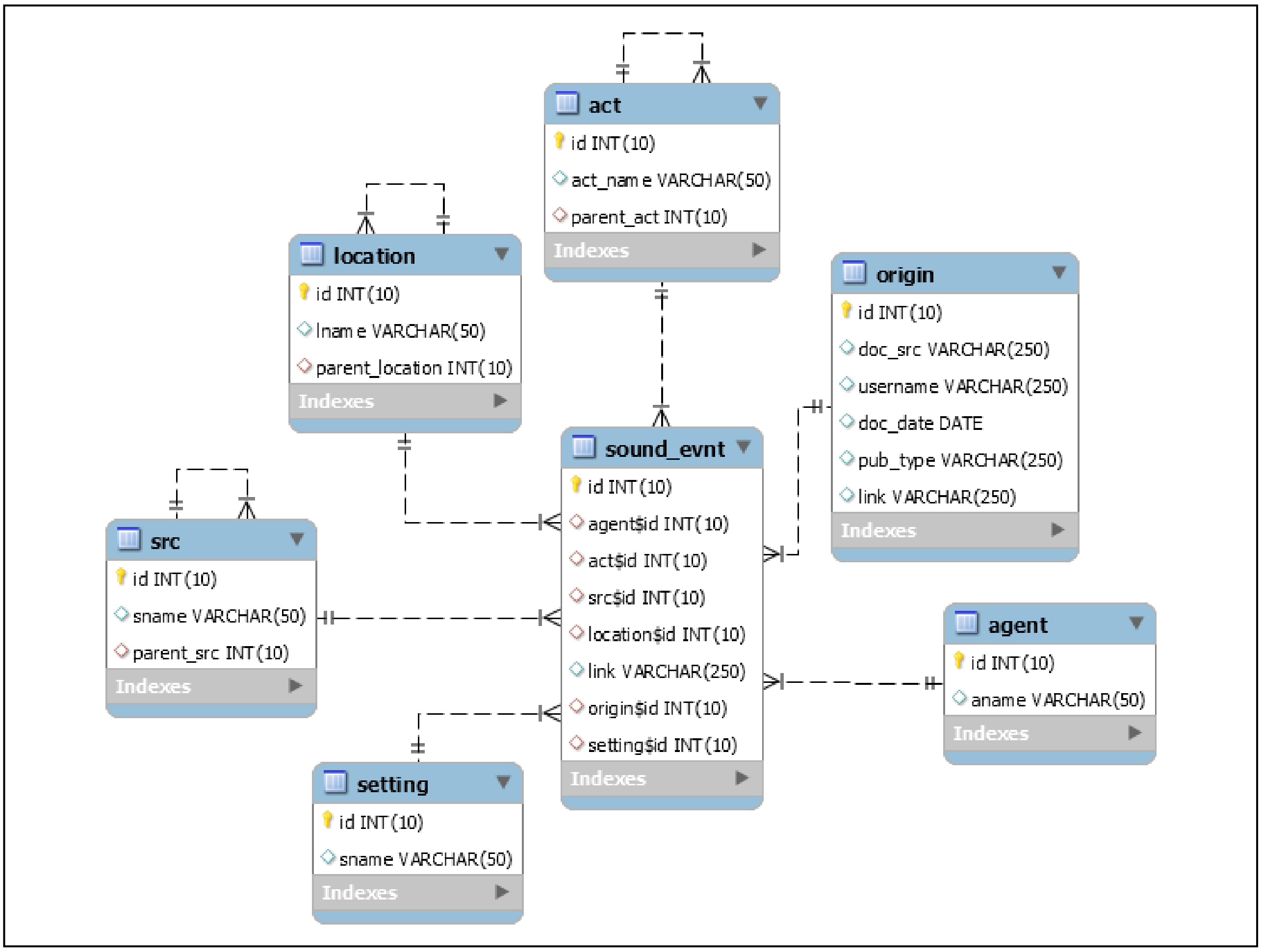


Figure 1: Database Entity-Relationship Diagram. The crow's feet indicate the direction of the one-to-many relationship. The primary keys of each table are represented by the yellow key icons. The foreign keys are identified through a combination of foreign table name and foreign field, in the format: *table_name$field_name*.

and Gjestland taxonomy [13]. Therefore, it is restricted largely to the following entries: urban, suburban, rural, wilderness and underwater.
The *setting* table is created separately from location for reasons that have more to do with database normalization than with the faceted taxonomy. Database design principles hold that redundant entries should not exist. For example, we could have sounds produced on the road, and thus our database will require a location entry 'road'. However, if location contains the setting as a top-level item, then we need to consider the urban road as distinct from the suburban road, causing there to be two or more entries for the same item, *road*.
The table labelled *origin* is added to the database to track the provenance of the sound events that are described. This includes previous empirical studies and online sound repositories.

### 2.2. Hierarchies within facets

A hierarchy of items can be defined for a specific facet. These hierarchies provide further structure to the content of the database, as well as defining relationships within these broad categories.
Three tables implement recursive relationships to provide the described hierarchical organization for

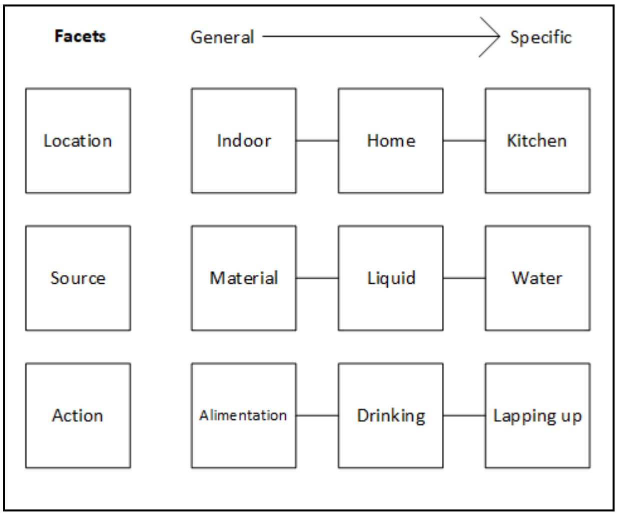


Figure 2: Example of hierarchies for location, source and action facets.

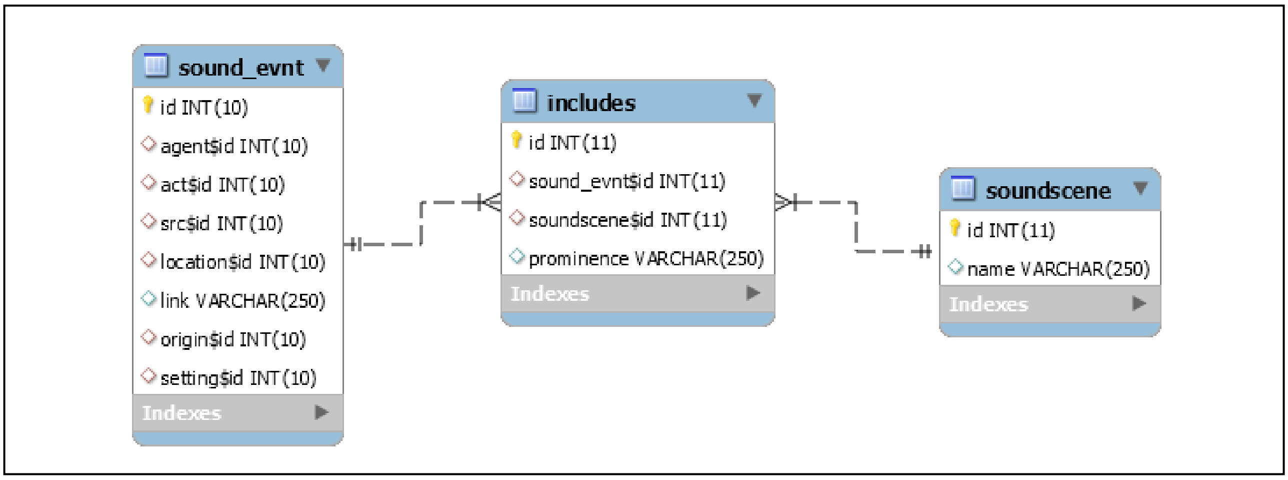


Figure 3: Modelling sound scenes in the database by grouping multiple sounds.

the facets of each sound. These tables are: source, action and location. For example, the sound of an engine can either be listed under engine or under vehicle. Through the use of a recursive relationship, we can define a hierarchical, broader-narrower term relationship between the engine and the vehicle. Thus, we can define a hierarchical relationship within each facet that provides a better description of the sounds.

In the early stages, the database will be validated using the taxonomies proposed by previous research into the classification of sounds. Eventually, the database will be further populated with sounds taken from the online repository, freesound.org. This site provides a large number of sounds with textual descriptions that can be mined to fill the database tables.

## 3. Discussion

A faceted taxonomy approach to sound classification has applications to research generally, but also practical uses for soundscape planning.

Soundscape research uses different conceptual approaches, making it difficult to directly compare the results that are obtained. A faceted taxonomy allows for common framework for the comparison of this research. For example, more research is necessary to understand why certain categorization principles are considered to be more important than others. The proposed taxonomy and accompanying database can be used to examine patterns in sound categorization.

The benefits to soundscape planners depend on the goals of the project. In general, the taxonomy can be used to compare the areas that are included in sound and noise documents. This provides the opportunity for an assessment toolkit for anyone working in this domain. For example, regulations and planning documents use a combination of sound sources and actions to carry out noise reduction strategies. These documents can be mined to extract details and then compared against the database to provide a categorical assessment of coverage and thoroughness.

In order to fully benefit from the implementation of this taxonomy, more work is required to add acoustical and hedonic dimensions. However, once these are fully implemented, the database could potentially be used to predict judgements of soundscapes and sound scenes.

## 4. Conclusions

There is evidence that people cross-categorize sounds according to a number of factors. Based on these results, researchers have proposed a number of classification systems for sounds, some of which use a combination of sound source and sound event. The hierarchical nature of these taxonomies makes the actual classification process challenging.

Given the limitations of hierarchical taxonomies, this paper proposes a faceted taxonomy that reflects the way people cross-categorize sounds. Four facets are used in the current taxonomy: agent, source, action, and location. Other facets that play a role in categorization include hedonic factors, familiarity, and acoustic descriptors. The current taxonomy does not consider these properties, and their addition is an important element in future phases. The taxonomy has been implemented using an open-source relational database management system (MySQL). A web interface will be implemented shortly to allow users to navigate the content and structure of the taxonomy.

## References

[1] J.G. Neuhoff: Introduction and history. In: J.G. Neuhoff, (ed.) Ecological Psychoacoustics, Brill, Leiden, 2004.

[2] J.A. Ballas: Common factors in the identification of an assortment of brief everyday sounds. J. Exp. Psychol. Hum. Percept. Perform. 19 (1993) 250–267.

[3] C. Guastavino: Categorization of environmental sounds. Can. J. Exp. Psychol. 61(1), 54–63 (2007).

[4] B. Gygi, G. R. Kidd, & C. S. Watson: Similarity and categorization of environmental sounds. *Perception & Psychophysics*, *69*, 2007 839–855.

[5] N.J. VanDerveer: Ecological acoustics: human perception of environmental sounds. Ph.D. thesis, Cornell University, 1979. Unpublished Dissertation

[6] F. Guyot, M. Castellengo, B. Fabre: Étude de la catégorisation d'un corpus de bruits domestiques. In: Dubois, D. (ed.) Catégorisation et cognition: de la perception au discours, Kimé, Paris, 1997.

[7] W.W. Gaver: What in the world do we hear?: an ecological approach to auditory event perception. Ecol. Psychol. **5**, 1993, 1–29.

[8] D. Dubois: Categories as acts of meaning: The case of categories in olfaction and audition. Cognitive Science Quarterly 1 (2000) 35–68.

[9] O. Houix, G. Lemaitre, N. Misdariis, P. Susini, & I. Urdapilleta: A lexical analysis of environmental sound categories. Journal of Experimental Psychology: Applied, 18 (2012) 52-80.

[10] C. Guastavino: Everyday Sound Categorization. In Computational Analysis of Sound Scenes and Events. Cham: Springer, 2018.

[11] C. Guastavino: The ideal urban soundscape: Investigating the sound quality of French cities. *Acta Acustica United with Acustica*, 92, 2006, 945-951.

[12] J. Salamon, C. Jacoby, and J.P. Bello: A dataset and taxonomy for urban sound research, Proc. 22nd ACM International Conference on Multimedia, 2014.

[13] A. L. Brown, J. Kang, & T. Gjestland: Towards standardization in soundscape preference assessment. *Applied Acoustics*, 72, 2011, 387-392.

[14] P. Lambe: Organising knowledge: Taxonomies, knowledge and organisational effectiveness. Chandos Publishing: Oxford, 2007.

[15] S. R. Ranganathan: Colon classification. Madras: Madras Library Association, 1957.

[16] L. Spiteri: A simplified model for facet analysis: Ranganathan 101. *Canadian Journal of Information and Library Science*, 23, 1998, 1-30.