

MESTRADO

MULTIMÉDIA - ESPECIALIZAÇÃO EM MÚSICA INTERATIVA E DESIGN DE SOM

TOWARDS A CROSS- DISCIPLINARY SOUND DESIGN METHODOLOGY: A FOCUS ON SEMIOTICS AND LINGUISTICS

Bruno Mascarenhas

September

2023

FACULDADES PARTICIPANTES:

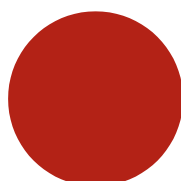
FACULDADE DE ENGENHARIA

FACULDADE DE BELAS ARTES

FACULDADE DE CIÊNCIAS

FACULDADE DE ECONOMIA

FACULDADE DE LETRAS



Abstract

This thesis is focused on the development of a methodology for sound design that encompasses auditory displays, psychoacoustics, semiotics, and linguistics. It explores the role of sound design in conveying messages (not just as an aesthetic or technical challenge, but as a form of communication), serving as an interface between users and devices, highlighting how sounds are communicative acts. It investigates the parallels between semiotics and sound design, interpreting how sounds can act as signs that represent something else, based on established social conventions and how we can use linguistics as a model to create non-speech sounds that, just like language, systematically convey meaning to the user. Through this comprehensive analysis, the work aims to contribute to the field by proposing a cross-disciplinary methodology for sound design, thereby enhancing the auditory experience in diverse contexts. Applying some of the linguistics principles (such as language components and illocutionary acts) to sound design (such as earcons) can result in more effective sounds for devices, providing a reference framework for audio communication. The goal is to contribute for a heuristic to describe and evaluate the production and optimization of systems of messages transmitted through sound.

Resumo

Esta tese está focada no desenvolvimento de uma metodologia para o design de som que engloba *auditory displays*, psicoacústica, semiótica e linguística. Explora o papel do design de som (não apenas como um desafio estético ou técnico, mas como forma de comunicação), na transmissão de mensagens, servindo de interface entre utilizadores e dispositivos e destaca como os sons são atos comunicativos. Também investiga os paralelos entre a semiótica e a linguística e o design de som, interpretando como os sons podem atuar como sinais que representam outra coisa, com base em convenções sociais estabelecidas, e como podemos utilizar a linguística como modelo para criar sons não-falados que, tal como a linguagem, transmitem sistematicamente significado ao utilizador. Através desta análise abrangente, o trabalho pretende contribuir para o campo, propondo uma metodologia interdisciplinar para o design de som, melhorando assim a experiência auditiva em diversos contextos. A aplicação de alguns dos princípios linguísticos (como as componentes da linguística e os atos ilocucionários) ao design de som (como os earcons) pode resultar em sons mais eficazes para dispositivos, fornecendo uma estrutura de referência para a comunicação áudio. O objetivo é contribuir para uma heurística que descreva e avalie a produção e otimização de sistemas de mensagens transmitidas através do som.

Table of contents

1. Introduction	9
2. Revision of literature.....	11
2.1 Auditory displays	12
2.1.1 Auditory displays and sonification	12
2.1.2 Designing auditory cues: Mapping musical parameters	13
2.2 Psychoacoustics and sound design.....	19
2.3 Semiotics.....	21
2.3.1 Road signs.....	25
2.3.2 Earcons.....	26
2.3.3 Linguistics.....	27
3. Methodology	31
3.1 Guidelines	32
3.2 Pragmatics for a hypothetical game.....	35
3.3 Sound Design model table	36

4. Practical cases.....	40
4.1 Simple sound systems	40
4.1.1 Cardiac Arrest Machine	41
4.1.2 Car Reverse Gear	42
4.1.3 Elevator.....	43
4.1.4 Traffic Light Crossing (Australian Crossings).....	43
4.2 Smartphones.....	44
4.3 Transformational earcons	47
5.Final considerations	49
5.1 Objectives satisfaction	49
5.2 Future work.....	50
References	52

Figures

Figure 1 - Some research fields relevant for sound design as communication	11
Figure 2 - General description of a communication system	12
Figure 3 - Umberto Eco's communication model	22
Figure 4 - Abstraction of product sounds based on its content	23
Figure 5 - Selection, organization, coordination, and structuring, for the items of perception and for the objects of experience	24
Figure 6 - Model of meaning for product sound design	24
Figure 7 - Major levels of linguistic structure	32
Figure 8 - Foundations in SD for Embedded Media: a Multidisciplinary Approach	47

Tables

Table 1 - Main auditory cue types, advantages, and disadvantages	14
Table 2 - Categorization of sound samples	23
Table 3 - Road sign system	25
Table 4 - Speech act table for hypothetical game	35
Table 5 - Model for User Interface Sound Design: 6 types and characteristics	37
Table 6 - Model for User Interface Sound Design: 6 types and 8 characteristics	38
Table 7 - Model for User Interface Sound Design: 3 types and 4 characteristics	38
Table 8 - Cardiac arrest	42
Table 9 - Reverse gear	42
Table 10 - Elevator status	43
Table 11 - Traffic light crossing	43
Table 12 - Smartphones sounds	44

1. Introduction

Context - Sound design (SD) is a discipline that emerged in the late 20th century, first acknowledged in the film industry in 1979, during the production of *Apocalypse Now* (Dienstfrey, 2016). Over the past few decades, the field has expanded beyond the confines of cinema and shows, permeating various sectors such as artistic and museum installations, video games, and even everyday products like cars, cell phones, and home appliances. The primary focus of SD is the transmission of messages through sounds, serving as an audio interface between users and devices. This enables the communication of information to users via audio, a concept that has seen growing sophistication in terms of knowledge and techniques. However, despite the advancements, there is a noticeable limitation - the lack of systematic methodologies for analyzing and creating sounds that communicate effectively, especially when compared with the vast amount of work in this area for visual signs. This gap in research motivates the pursuit of this study focused on the development of a methodology for sound design based on semiotics. It is focused on earcons - abstract synthetic tones - that can be used to create structured and auditory messages from a set of simple rules (since they are arbitrary and any signified can be represented), mainly the one-element type (as the sounds produced for cardiac arrest machines, car reverse gears, elevators, traffic light crossings) but also the transformative (as some of the mobile phones sounds).

Problem, hypothesis, and objectives - The growing use of electronic devices and environments in daily life led to a growing necessity of consistent and effective user interface sound design methodology that contributes to a better overall experience (user experience). Yet, in contrast with the abundant research concerning the field of visual user interface design, audio user interface design research concerning methodologies for creating and evaluating sounds is still scarce. So, this research pursues the following question: *how can we develop a methodology to create and evaluate sounds that communicate effectively?* The significance of this study lies in its potential contribution to the field of SD. It aims to explore the role of sound in communication providing a systematic methodology for effectiveness by applying semiotics principles to sound design. The goal is to develop a conceptual tool for sound designers, enhancing the way abstract sounds are used for communication. It approaches sound design not just as an aesthetic or technical challenge, but as a form of communication, with its own set of

rules and principles and highlights sound designers as communication engineers and how sounds function as communicative acts.

Methodology - This dissertation proposes a sound design methodology based on the findings of semiotics focused on linguistics, particularly the pragmatic component and its speech acts. This choice aligns with the exploration of analogous concepts in the visual domain, where semiotics has made substantial contributions, paralleling its applications to advance our understanding of sound design. Isomorphisms allow human cognition to transcribe an unsolvable problem from one domain to a different domain where the problem might be more easily addressed. Therefore, a methodology for sound design is proposed by establishing relationships between linguistics components and audio signals, associating the two domains. By applying some of the linguistics principles (such as language components and illocutionary acts) to sound design (such as earcons) it aims to contribute to more effective sounds for devices, providing a reference framework for audio communication¹.

In the guidelines chapter, we define how we can apply the field of linguistics to sound design and use it to create guidelines for conveying messages with abstract sounds. In the pragmatics for a hypothetical game chapter, we delve into how we can use Searle's speech act theory to classify the different messages created by the abstract sound systems. In the sound design model chapter, we address how to organize information.

The goal is to contribute for a heuristic that intends, exactly, to describe and evaluate, how to produce and optimize, systems of messages transmitted by sound.

Document Structure - This document is divided into introduction, revision of literature methodology, practical cases, and final considerations, each one of these addressing a specific aspect of the research. The introduction provides a detailed context of the study, followed by the research question, objectives, and proposed methodology. The revision of literature delves into the existing research on methodologies for sound design and evaluation in fields such as auditory displays, psychoacoustics, and semiotics. It also explores the role of linguistics in sound design, approaching the application of Searle's pragmatics to sound design. The methodology details the development of a comprehensive system for creating communicative sounds based on linguistics. It includes the creation of model tables and guidelines. The practical cases chapter applies the developed methodology to several cases, such as simple sound systems, smartphones, and transformative earcons, aiming to evaluate its effectiveness and reliability. The final considerations discuss the results and their implications for the field of sound design.

¹ For instance, a directive sound in a user interface could be used and understood to instruct the user to perform a certain action, such as saving the work before the system shuts down. An expressive sound, on the other hand, could convey the system's status or reaction to the user's actions, such as a success sound indicating that a task has been completed successfully.

2. Revision of literature

This revision of literature aims to examine different research fields that influence sound messages communication such as auditory displays, psychoacoustics, and semiotics as shown in Figure 1. By exploring this body of knowledge, this revision will provide a comprehensive understanding of the current state of research in SD and lay the grounds for the development of a methodology to create and evaluate sounds.

Disciplines	Sound Design as communication
Auditory displays	<ul style="list-style-type: none"> • Use of sound to communicate information about the state of a computing device to a user • psychology, computer science, and audio engineering, etc.
Psychoacoustics	<ul style="list-style-type: none"> • How our brains process and interpret the information carried by sound waves • physics, psychology, cultural studies, etc.
Semiotics	<ul style="list-style-type: none"> • Sound as a cultural process and a system of communication. The study of signs (something that stands in for something else) • philosophy, linguistics, sociology, psychology, anthropology, etc.

Figure 1 - Some research fields relevant for sound design as communication

First, the focus will be on auditory displays, mainly sonification and auditory cues. It includes an analysis of the use of auditory displays and the importance of sonification in transmitting information effectively, as well as mapping parameters for designing auditory cues.

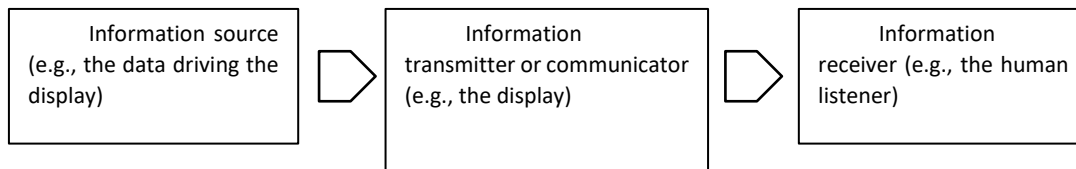
Secondly, the focus will shift to psychoacoustics and sound design, examining objective sound perception and the subjective parameters that influence sound design.

Finally, the focus will turn to semiotics and Searl's linguistics, presenting road signs as a prominent visual example of a semiotic code and exploring the use of auditory cues as symbols to convey meaning to sounds.

2.1 Auditory displays

Auditory displays use sound to communicate information (Walker & Nees, 2011) and offer a relay between the information source and the information receiver as in any information system. In this case, the data of interest are conveyed to the human listener through sound, as shown in Figure 2 (Kramer, 1994; Shannon, 1998/1949).

Figure 2 - General description of a communication system



2.1.1 Auditory displays and sonification

The field of auditory displays encompasses the use of sound to communicate information about the state of a computing device to a user, with sonification being a subtype of auditory displays that specifically use non-speech audio to represent data relationships. Sonification involves turning data into recognizable relationships within an audio signal for the purpose of better communication or interpretation (Kramer et al., 1999). This multidisciplinary approach draws upon diverse fields such as psychology, computer science, and audio engineering. It can best be characterized as an amalgamation of important insights from multiple disciplines, as pointed out by Kramer (1994).

The 1999 collaborative Sonification Report (Kramer et al., 1999) outlined four key areas that require attention in the theoretical description of sonification. These include 1) taxonomic descriptions of sonification techniques based on psychological principles or display applications; 2) descriptions of the types of data and user tasks suitable for sonification; 3) a discussion of the mapping of data to acoustic signals; 4) an examination of the factors limiting the use of sonification.

According to McGee (2009), in sonification, designers consider factors such as perception, cognition, acoustics, and psychoacoustics. The author identifies eight fundamental characteristics of sound waveforms: frequency, amplitude, phase, envelope, spectrum, shape, velocity, and wavelength, which can be used to map data to sound. Additionally, he refers, data can also be mapped to synthesis techniques, such as amplitude modulation, frequency modulation, additive synthesis, subtractive synthesis, and granular synthesis.

Spatial auditory displays offer another dimension to the mapping process, shifting the focus from "what does this data sound like?" to "where is this data in space?" In musical sonification,

musical parameters, such as tempo, rhythm, time signature, and tuning, can be used to add aesthetic appeal to the sonification as pointed out by McGee (2009).

The author also highlights that loudness, masking, and Doppler shift are important perceptual issues of psychoacoustics that must be considered in the mapping process. He concludes that there is no technical basis for mapping data to sound and that it is up to system designers to determine the appropriate mapping process for a given application (McGee, 2009).

Auditory altimeter, clocks with audible ticks and chimes, and storm and weather sonification are examples of sonification following presented.

- Auditory altimeter - A device used in skydiving that provides altitude information through sound. It typically emits a series of beeps or tones that vary in frequency or tempo depending on the skydiver's altitude above ground level. Application - skydivers rely on auditory altimeters to monitor their altitude during a jump. The changing pitch or rhythm of the auditory signals helps them gauge their altitude without having to look at a visual display, allowing them to focus on their freefall and parachute deployment.
- Clocks with audible ticks and chimes - These clocks are designed to produce audible sounds with each passing second and distinctive chimes at regular intervals (e.g., every 15 minutes). The tick-tock sound and chimes provide temporal information. Application - Auditory clocks are often used in various settings, including homes and offices. They serve as timekeeping devices that help people keep track of time without needing to constantly look at the clock face. The audible cues can be especially helpful for individuals with visual impairments.
- Storm and weather sonification - Sonification is used in meteorology to translate weather data into audible representations. Various meteorological parameters, such as temperature, wind speed, and atmospheric pressure, are mapped to sound properties like pitch, volume, and rhythm. Application: Weather sonification can assist meteorologists in analyzing complex weather patterns and trends. By listening to the auditory representation of weather data, meteorologists can detect changes and anomalies in real-time. It can also be used as an educational tool to help the general public understand weather conditions better, especially for those who are visually impaired.

2.1.2 Designing auditory cues: Mapping musical parameters

Auditory cues are a potential way of encoding information in an auditory display (Shinn-Cunningham et al., 1997). They are like visual icons in graphical user interfaces, serving the purpose of conveying key information to the user in a quick and effective manner. The forms of

auditory cues and their applications are diverse, and designers must make informed decisions about which form to use in each situation based on its benefits and drawbacks.

The types of auditory cues are described in Table 1 (McGookin, 2019).

Table 1 - Main auditory cue types, advantages, and disadvantages

<i>Cue type</i>	<i>Description</i>	<i>Advantages</i>	<i>Disadvantages</i>
Earcons	Short-structured music-like sounds that can be manipulated using music rules to encode information.	<ul style="list-style-type: none"> • Relatively rich messages can be generated from a set of simple rules. • As the encoding between sound and what it signifies is arbitrary, any signified can be represented. • Earcons have been extensively evaluated, and there are good guidelines for their creation and use. 	<ul style="list-style-type: none"> • As the encoding between sound and signified is arbitrary it must be learned, requiring training. • The need to have sounds be perceptually different can limit the size of a set of possible earcons.
Auditory icons	Everyday sounds used to represent computer events by analogy of everyday sound-producing events.	<ul style="list-style-type: none"> • The relationship between sound and what it represents is based on common knowledge, so training is not required. • Auditory icons can employ more dynamic elements, such as physics simulations. • Guidelines are available to support effective auditory icon creation. 	<ul style="list-style-type: none"> • It can be hard to identify sounds that map to all needed events. • Manipulating sound across real-world parameters is nontrivial. • Auditory icons assume domain knowledge of the everyday use of sounds, it is not clear if this will work cross-culturally.
Spearcons	Speech significantly speeded up to sound like an auditory texture.	<ul style="list-style-type: none"> • Can easily create spearcons for required signifieds. • Sounds can be trivially generated. • Can be used to overcome the “brittle” nature of other. 	<ul style="list-style-type: none"> • Evaluation of spearcons has largely been focused on menu structures, it is not clear how spearcons will perform in many of the situations auditory icons and earcons work.
Spindex	A modification to spearcons to emphasize the first letter and provide quicker navigation through alphabetized menus.	<ul style="list-style-type: none"> • Can improve the access speed in alphabetized menu structures. 	<ul style="list-style-type: none"> • Is limited to supporting alphabetized menus and is not a general-purpose technique (as with auditory icons, earcons and potentially spearcons). • A spoken (or visual) menu item is still required. Spindex only provides a faster way.
Musicons	Small snippets of music used to represent information associated with the music or lyrics.	<ul style="list-style-type: none"> • Supports individuals to create their own link between the musicon and referend. • Can potentially provide a degree of privacy in 	<ul style="list-style-type: none"> • There is limited study and guidelines on selecting effective musicons.

Cue type	Description	Advantages	Disadvantages
		the meaning of the musicon.	
Source: McGookin (2019, p.249).			

It examines various cue types employed within sound design. Each is presented with a concise description of its characteristics, followed by a balanced assessment of its advantages and disadvantages. It aids in developing a clear understanding of the distinctive attributes of each cue type. For professionals in the field, the table can assist in making informed choices when selecting an appropriate cue type for specific projects or applications, providing a structured overview of the pros and cons, enabling sound designers to align their decisions with project-specific needs.

By examining diverse cue types, the table underscores the interdisciplinary nature of sound design where fields such as semiotics, linguistics, and psychoacoustics converge in the development of effective auditory communication methods.

2.1.3 Earcons

For this study, the focus is earcons. As first proposed by Blattner et al. (1989), earcons are abstract, synthetic tones that can be used to create structured auditory messages. According to the authors, they can be constructed using musical principles and employ musical motives as their core form. Musical motives, defined as brief successions of pitches arranged to produce a rhythmic and tonal pattern that is sufficiently distinct to function as an individual, recognizable entity, have been used in musical compositions and advertising for representation purposes. As referred in Table 1, although requiring learning and different sounds, earcons can generate rich messages from a set of simple rules. Another advantage is that any signified can be represented since the encoding between sound and what it signifies is arbitrary. Also, there are reference guidelines for their creation and use. Earcons were chosen as the focus of this study primarily due to their abstract nature.

Brewster et al (1994, 1995b) carried out significant work to determine the design of effective earcons and their benefits in a user interface. They found that earcons should be created with a musical approach² and that providing maximum variation between auditory parameters reduces the likelihood of misidentification. Leplâtre and Brewster (2000) later noticed that the use of subtle musical elements could increase the number of earcons.

The knowledge derived from icon design theories can be applied to formulate equally successful strategies for designing earcons. Just as icons can be categorized, earcons are also classified into three groups: representational, abstract, and semi-abstract (Blattner et al., 1989).

² By using musical timbre and complex rhythmic pitch structures.

Representational earcons, analogous to pictorial icons, can be created through various schemes. For instance, digitizing natural sounds from the surrounding environment is one approach, mirroring the use of highly representational images for icons. These earcons can serve as audio messages, with similar advantages and disadvantages as pictorial icons. An investigation by Gaver (1986) delved into representational earcons or auditory icons, which caricature real-world sounds like bumps, scrapes, or files hitting mailboxes. Gaver identified three types of mappings between data and auditory representation: symbolic, nomic, and metaphorical.

Symbolic mappings rely on social conventions for meaning, akin to using applause to indicate approval. Nomic representations are physical, such as using the sound of a closing metal cabinet to represent closing a file. Metaphorical mappings, drawing on similarities, could involve a falling pitch to signify a falling object. Gaver emphasized that auditory icons need not be exact replicas of the objects they represent but should capture essential features, similar to representational icons.

Abstract earcons follow a method similar to Marcus's (1984) concept of "elements" geometric shapes and marks that serve as foundational components for creating sets of icons. In the earcon design approach advocated here, single pitches or pitch groups function as the building blocks. These building blocks, called motives, are sequences of pitches that create distinct audio patterns often characterized by rhythmic simplicity and pitch design. The brevity and distinctiveness of motives make them useful tools for composing earcons. Complex objects can be expressed using compound earcons composed of these motives or single pitches. This approach offers systematicity, modularity, and the ability to transform, combine, and inherit motives to create related families of earcons. Grouping motives into families allows for differentiation in sound among earcons with similar meanings. This design approach is practical for contemporary computer systems, eliminating the need for specialized equipment.

McGookin and Brewster (2004a, 2004b) expanded on Brewster's work and used a family of transformational earcons to investigate concurrent presentation. Walker and Kramer (1996) distinguished between categorical and linear sound dimensions and proposed mapping data parameters accordingly. For example, McGookin and Brewster (2004b) used the tempo of a melody to represent the intensity of a theme park ride and the register to represent its cost.

Brewster et al. (1995a) summarized guidelines for earcon design categorized by musical parameters - timbre, rhythm and pitch, register and spatial location - as detailed next.

Timbre

Earcons with musical timbres tend to be more effective than those with simple sinusoidal tones. To maximize effectiveness, it is important to use timbres that are easy to differentiate. When playing earcons concurrently, different timbres from the same instrument family can help prevent confusion. Earcon identification was improved by using different timbres from the same musical family (McGookin and Brewster 2004b).

Rhythm and Pitch

Effective earcons rely on rhythm and pitch for discrimination. The pitch structure of earcons can also help with metaphorical mappings of data. Earcons should be short, with a maximum of six notes, and have different tempos for differentiation. Brewster et al. (1995a) noted that earcons with up to six notes that are as short as 0.03 seconds can be effective.

Register

When making absolute judgements of pitch, registers should not be used. If register is utilized to represent a data parameter, it should only be a few, with large differences of several octaves. McGookin and Brewster (2004b) found that earcons with register information (use of different octaves) were identified with around 70% accuracy of recognition but removing the need to recall register increased accuracy to around 90%.

Spatial Location

Spatializing earcons from different families can enhance their presentation and reduce the time it takes to present compound earcons. Concurrent spatial presentation of earcons from the same family can improve performance, but McGookin and Brewster (2004a) note that the improvement may not be enough to support effective earcon identification.

2.1.3.1 Types of earcons

Blattner et al. (1989) identified and detailed four types of earcons: one-element, compound, inherited and transformed.

- **One-element earcons** - The most basic form of earcons (the one we will be focusing on mostly) is "one-element earcons", which represents a single mapping between one motive and the information it represents. Examples of one-element earcons include sounds used to represent text messages, new emails, and calendar reminders on a smartphone. This category is rooted in the concept of utilizing a single audio element to convey a specific meaning or function, drawing an analogy to linguistic components like words and letters. One-element earcons show a remarkable versatility by incorporating digitized sounds, synthesized audio, single notes, or motives. This approach allows for an intricate play of combining these elements, akin to linguistic amalgamation, resulting in nuanced auditory messages. The one-element earcons can be further sub-divided into the following subtypes:
 - Single-pitch earcons - These earcons rely on a singular note with attributes encompassing pitch, duration, and dynamics. Their simplicity makes them particularly apt for representing straightforward interface actions like key clicks, cursor movements, or selection mechanisms. Additionally, they can adopt metaphorical meanings; for instance, the use of a pitch gradually decreasing in

volume may symbolize the act of a computer shutting down, creating a meaningful audio metaphor for the application's function.

- Single-motive earcons - This subtype exhibits a heightened complexity, incorporating a broader array of attributes including rhythm, pitch, timbre, register, and dynamics. As a result, single-motive earcons can represent more intricate interface elements, such as conveying error messages, system information, windows, and files. Their versatile attributes lend themselves to nuanced representation, where the combination of attributes can indicate the nature or location of the represented interface element.
- **Compound earcons** - Moving beyond the confines of single-element compositions, compound earcons delve into the realm of employing multiple audio elements to articulate intricate meanings or functions. This approach hinges on three core principles: combining, inheriting, and transforming. By intertwining these principles, compound earcons can be forged, which, much like the approach to visual icons, enables the representation of diverse system functions or entities. In essence, these compound earcons are composed through the judicious selection and juxtaposition of audio elements, leading to the creation of a comprehensive set of auditory cues. This systematic approach bears resemblance to the use of repeated visual elements in icons, streamlining the user's capacity for identification and retention. Through the amalgamation of these auditory elements, a rich tapestry of auditory cues can be woven, enhancing the user experience, and fostering a more intuitive interaction with the digital environment. For example, combining individual earcons that are linked to 'open' and 'save' for example creates compound earcons, indicating a sequence of actions, in this case opening a file followed by saving it.
- **Inherited earcons** - The concept of inherited earcons pertains to a network of earcons linked through shared or modified attributes, thereby forming what's referred to as "family earcons." Hierarchical families are examined, wherein each earcon operates as a node within a tree-like structure, with earcons on each level consisting of a sequence of a specific number of elements. This structure involves the modification of attributes at each level, encompassing rhythm, pitch, timbre, dynamics, and register. The following rules govern the creation of hierarchical families of earcons:
 - Identification of entity families and their related attributes.
 - Assigning distinct rhythmic structures to each family.
 - The family rhythm, akin to a last name, becomes the foundation of the earcon representing the highest-level entity.
 - Entities at the second level are conveyed by a two-part earcon: the family rhythm followed by pitches rhythmically aligned.
 - Third-level entities employ a three-part earcon, combining the family rhythm, level-two motive, and a slightly altered version of the level-two motive with different timbre and pitch.

For hierarchies extending beyond three levels, attributes like register and dynamics are employed. These family earcons convey layered information, where the family rhythm denotes the entity family, the second part signifies a specific type within that family, and the third part identifies precise details like an error message. While the given structure is advantageous, alternative approaches might involve sequences of less closely related elements, creating complex meanings when concatenate, as exemplified by the combination of fire and water sounds.

- **Transformational earcons** - Transformational or transformed earcons involve modifying a musical segment by altering its attributes. These transformations, aiming to preserve perceptual equivalences, are analogous to the concept of preserving musical shapes or contours. The transformations, such as inversion, retrogression, or retrograde inversion, can significantly modify the identity of the musical segment. These ideas draw parallels with Arnold Schoenberg's (1951) musical transformations, which reflect the original tone row across different axes. However, the question of whether such transformations genuinely yield perceptual equivalences remains a topic of debate, emphasizing the complexity of the auditory perception involved.

2.2 Psychoacoustics and sound design

Psychoacoustics is the study of the relationship between the physical properties of sound waves and our perception of sound. It seeks to understand how our brains process and interpret the information carried by sound waves. This field encompasses a range of disciplines, including physics, psychology, and cultural studies. By understanding psychoacoustics, sound designers and engineers can create sounds that have the desired emotional, cultural, and perceptual effects.

One of the main distinctions in psychoacoustics relevant for this work is the difference between an objective perception of sound properties and a more subjective or emotional perception of sounds. We will be looking into studies and psychoacoustics parameters that will be important for the development of our table in the methodological chapter.

There are many psychoacoustics phenomena that can be important to consider when developing sounds for certain contexts such as the: threshold of hearing, localization, interaural time difference, interaural intensity difference, perception of loudness, loudness scales and weighing, the doppler effect and many others. However, a more in-depth analysis of these phenomena goes beyond the scope of this research.

Some of the most important studies in this field have been on auditory warnings, starting in the 1980s with Patterson's work (1982) which examined how users emotionally responded to alarms in civil aircraft. The study concentrated on designing effective auditory warnings, introducing key design principles. The research addressed issues with warnings, presented a

model for immediate-action warnings, and conducted testing. The study revealed confusion arising from similarities in temporal characteristics, particularly repetition rate. Listening tests and clustering analysis supported this observation. To minimize confusion, specific temporal and spectral attributes for flight-deck warning sounds were defined. Urgent warnings required distinctive patterns with short intervals between pulses, while non-urgent warnings needed longer intervals. Spectrally, components in specific frequency ranges were identified, with suggestions for harmonic components and fundamental frequencies. Additional elements were recommended for immediate-action warnings.

Kuwano et al. (2000) explored the impact of frequency components and off-time on danger perception. Their investigation indicated that higher frequency sounds heightened the sense of danger across various locations, except for wide frequency range stimuli. Shorter off times also intensified danger perception. The study proposed using audible warning signals with wide frequency ranges to effectively convey danger messages. Such signals would be less affected by background noise and accessible to individuals with hearing impairments, making them desirable for auditory warnings.

These are important references in establishing guidelines to justify the use of certain sound properties to achieve specific effects and contributed to guidelines that are still used today in alarm systems and other applications.

Though, while psychoacoustics provides valuable insights into the perception and emotional effects of sound, it does not provide a comprehensive methodology for sonification. It is difficult to support decisions for all sound types with nothing but verified phenomenology based on psychoacoustics research.

As mentioned before, there is no established method for converting data into sound and it is up to the designers of the system to decide on the best approach for their specific use case (McGee, 2009).

Garner et al. (2010), and Lopes et al. (2017), conducted research into the influence of loud and high-pitched sounds on evoking feelings of fear. The latter study (2017) specifically examined how various pitches and loudness levels of sounds impact the emotional states of individuals in terms of tension, arousal, and valence dimensions. In a related study, Toprac and Abdel (2010) investigated emotions such as fear, suspense, and anxiety by considering three key attributes of sound effects: volume, timing, and source. Their findings showed that a sound design featuring high volume and synchronized timing with visual cues is most effective at inducing fear, with the source of sound effects also playing a contributing role.

Salselas et al. (2020) investigates the role of sound design in immersive audiovisual environments. The authors observe that sound design, which encompasses elements such as music, audio effects, and foley, has emerged as a potent tool to navigate the challenges presented by these environments. It possesses the ability to manipulate and intensify visuals, thereby

assuming a fundamental role in storytelling. By creating the illusion of "being there," sound design situates the viewer in a specific point of view and guides them through a particular narrative. In this context, sound design becomes instrumental in directing the user's attention and emotions.

The authors also emphasize the concept of consistency as pivotal in this context. It can be perceived as a salient event that captures and drives attention, achieved by matching visual and auditory *stimuli* in terms of synchronism and prominence. For instance, the redundancy of auditory and visual information can strengthen the sensory signal in a disruptive, uncorrelated sensory environment. Salselas et al. (2020) further note that consistency also facilitates emotional engagement. This is achieved by matching visual and auditory acoustic information in terms of visual space and acoustic spatialization, such as reverberation, delay, and sound spatialization. This can lead to an acoustic experience that drives emotional states.

The research concludes that musical *stimuli* may be remarkably effective in driving emotional states and engaging motivation and attentional resources. A stimulus that is highly emotionally relevant will naturally engage attentional resources. Therefore, in the context of immersive audiovisual environments where the user has agency, sound design can be a powerful tool for guiding the user's attention and emotions, thereby enhancing the narrative experience (Salselas et al., 2020).

2.3 Semiotics

In this section, we will address how sounds can act as signs that represent something else based on established social conventions and how we can use linguistics as a model to create non-speech audio that, just like language, systematically convey meaning to the user.

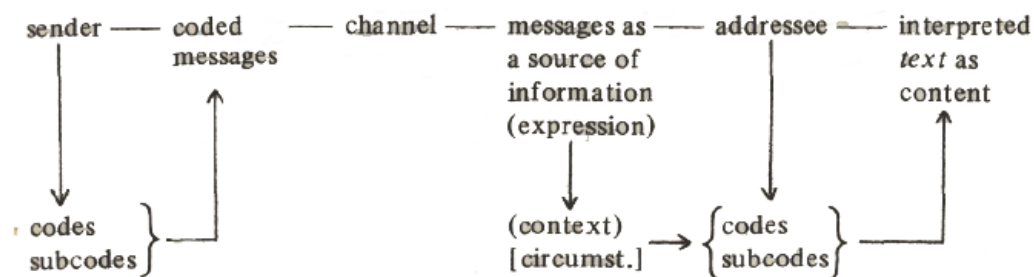
Authors in the field of SD have mostly made use of semiotics and the tripartite division of the sign, as well as its classification into index, icon and symbol, as we will see. However, if one observes the task of codifying, through sound, a set of messages that serve as an interface between a user and a device, one realizes that the question of whether the sound used has a relation of similarity (icon) or is a pure convention (symbol), or even, a matter dealt in psychoacoustics, whether it already brings with it a connotation of danger or not, is not fundamental (except when it comes to the ease with which users learn the code). More important to our work is to notice that something is being transmitted to the user and that he should try to interpret and act accordingly (especially if this means danger to himself, others, or the device itself). The essence of a sound designer is, first and foremost, to encode a message using sounds that work (often in partnership with visual or other signifiers). In this sense, we must be aware of the elementary structure of communication.

Eco (2000) defined semiology as the general theory of communicative phenomena, seen as the elaboration of messages based on codes agreed upon as systems of signs, and where isolated systems of signs are classified as “semiotic” if they are formalized or can be formalized. He defined as sign “everything that, on the basis of a previously accepted social convention, can be understood as something that is in the place of another thing” (p.11). The author arrives at this proposal after identifying the two classic definitions provided by the pioneers of contemporary semiotics: Peirce and Saussure. Saussure coined the term “semiology” in his *Course in General Linguistics* (1995) having been the privileged term in Europe. In the Anglo-Saxon world, Peirce’s designation “semiotics” was preferred (1999).

In addition to words and language, Eco addressed non-linguistic and even natural signs, which signify, based on a code, or previous learning, grounding, for instance, semiotic approaches of sound design. So, signs may take the form of words, images, sounds, odors, flavors, acts, or objects. But such things have no intrinsic meaning and become signs only when we invest them with meaning.

Eco rewrote the communication model (Figure 3) and stated that "the information of the message is only reduced by the addressee when he selects a definite interpretation" (1976, p.141).

Figure 3 - Umberto Eco’s communication model

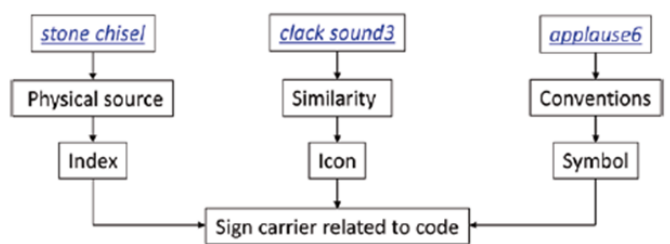


A sender (not necessarily a person; it may be nature, a traffic sign, a Morse machine, a wireless speaker, etc.) sends a message, transmitted by means of a code through a channel (e.g., sound waves). The addressee (could be a listener) is a sort of interpreter who uses his knowledge of the same code to discern the meaning of the message.

As product and/or interface sound design focuses essentially on the transmission of messages with sounds, researchers in this field (Mcgookin, 2020; Suied et al., 2005; Jekosch, 2005) have turned to semiotics to organize and guide their descriptions and methodologies, mainly Peirce’s (1839-1914) semiotics and the tripartite division of the sign, as well as its classification into index, icon and symbol (1999). Peirce's semiotics is rooted in his philosophical approach of pragmatism, which emphasizes the practical implications of knowledge and the role of human action in the formation of meaning.

The studies in this field tend to focus more on the question of the advantages and disadvantages of the relationship established between the sound (as signifier) and its referent (as constituents of the message that will be interpreted) being based on similarity (icon, auditory earcon) or on convention (symbol, coined generically as earcon) and sometimes also on causality (index). For example, Chandran et al. (2015) stated that understanding of the semiotic relations of sounds and product interface controls can support product designers in assigning sounds to products in a systematic and innovative manner. The authors coded 670 sounds based on the three referred categories of abstractions, as shown in Figure 4 and Table 2 (Chandran et al., 2015).

Figure 4 - Abstraction of product sounds based on its content



Sound samples	Index	Icon	Symbol
Chipmunk chatter		✓	✓
Cannon fires	✓		✓
Water gushing	✓		
Winding down theme		✓	
Cola burp	✓		✓
Cash register rings 2	✓		
Child toy duck quacks		✓	✓
Office phone ring	✓		✓
Chop vegetables	✓		

Table 2 - Categorization of sound samples

Studies on human-device sound interaction should consider not only the immediate aspects of the interaction, but also aspects underlying the cultural and social context in which it takes place. Semiotics applied to software or devices sound design can fill this gap. The motivation for the use of semiotics in interface sound design rests as well on the fact that, although sound interface elements may be conceived as “tools” in analogy to the real ones, they do not exist as “physical” objects, but as signs.

We perceive fractions of the world through our sensory organs, then treat them as sign carriers. Lastly, we associate objects with them and, consequently, meaning evolves. “Semiosis” takes place (Jekosch, 2015). Based on the model of semiosis as shown in Figure 5, Jekosch proposes a model of meaning for product sound design, as expressed in Figure 6.

Figure 5 - Selection, organization, coordination, and structuring, for the items of perception and for the objects of experience

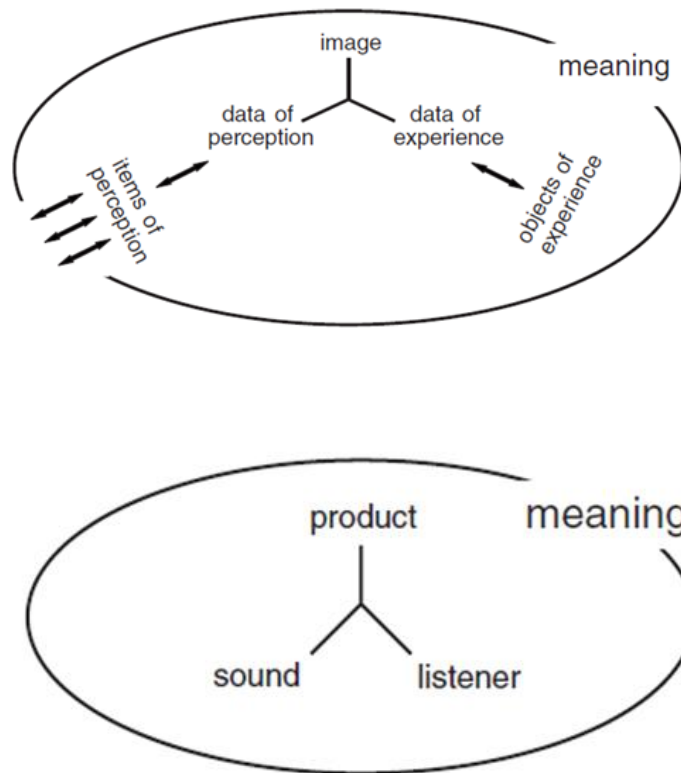


Figure 6 - Model of meaning for product sound design

Product use is a communicative act where the sound is one possible carrier of information regarding the product. When using the product, semiosis takes place in the listening user. Data of perception are, then, correlated with data of experience, and, finally, end meaning is associated with the auditory event specifically, and with the product generally. All this happens in a communicative context (p. 206).

A symbolic system is defined as a structured set of signs (linguistic, visual, auditory, or other) that possess relationships of contrast and/or combination with each other. Two semiotic systems are detailed next: visual signs (more specifically road signs), and auditory signals (more specifically earcons). We will see that as the first uses colors and shapes to contrast, the second uses variations in sound patterns and rhythms to differentiate.

2.3.1 Road signs

Traffic signals are an example of a semiotic system that can convey a lot of information through few means. Table 3 (Wagner, 2006) illustrates how the road sign system works.

Table 3 - Road sign system

	Danger warning signs	Regulatory signs	Informative signs
Sign-perception	To warn and inform road users of a danger	Obligations, restrictions or prohibitions	Travel information (direction, distance, place)
Colour semiotics	Yellow or Red	Red or blue	Green or blue
Shape semiotics	Diamond shaped or triangle	Circular, octagonal, triangle inverted	Rectangular, or square

Source: Wagner (2006)

According to Saussure (1995), the focus of his structuralist perspective on semiotics was on the opposing differences between signs. He believed that the crucial connections in structural analysis are binary contrasts such as nature vs. culture and life vs. death. So, he argued that the idea of a concept is not determined by its own qualities but rather by contrasting it with other concepts in the same system. In other words, the definition of a concept is formed by its relationship to other concepts in the system.

As Table 3 illustrates, the road sign system is a prime example of how contrast plays an important role in semiotics. In this symbolic system, signs are designed to convey meaning to drivers and other road users with colors, shapes, and images.

For example, yellow is used for warning signs, and green for directional signs. This differentiation by color creates a clear contrast between different types of signs, allowing drivers to quickly identify their meaning.

The shapes of the signs are also designed to create contrast, making it easier for drivers to distinguish between them. For instance, regulatory signs are circular (plus the specific cases of the shapes octagonal - for stop; and triangle inverted - for giving priority - that unequivocally permit drivers to identify these essential signs from the back). Warning signs are triangular or diamond shaped, and informative signs are rectangular or square. This distinction allows drivers to immediately identify the type of sign they are encountering, even from a distance.

In traffic signs, there is no phenomenological reason that justifies all the choices of sign attributes. The system is based on the idea that differences in sign features, like color or shape, make them easy for drivers to see and remember.

This contrasts with a phenomenological approach, which would try to justify each attribute based on previous empirical evidence.

According to Peirce (1999), there are three broad categories of relationships between signifier and signified: iconic, symbolic, and indexical. Iconic relationships refer to when the signifier visually resembles the signified. For example, the illustration on a pedestrian crossing sign looks like a person crossing the street. The resemblance is so clear that even without knowing the purpose of the sign, an individual could guess the meaning. Symbolic relationships, on the other hand, are arbitrary, with no assumption that the individual will understand the relationship between the signifier and signified without being told. For instance, in road sign design, the UK road sign indicates the national speed limit, but the relationship between the signifier (the road sign) and the signified (the speed limit) is not immediately obvious. Finally, indexical relationships are less clear and have a causal relationship between the signifier and signified. For example, smoke is a signifier of fire, as the fire causes the smoke. In road sign design, understanding the relationship between the signifier and signified is crucial for effective communication and road safety. The mapping of a sign to a category may change over time and can also be culturally situated. It is important to consider that cues designed within one cultural context may not work in another.

2.3.2 Earcons

Mcgookin (2019) makes an analysis of what each auditory cue represents in a semiotic system. As mentioned before, for our purpose, we will be focusing on earcons. In semiotics, earcons are considered to lie towards the symbolic end of the mapping continuum, meaning that they have an arbitrary relationship between the sound and its meaning. This provides a level of privacy or security to the sound and its representation, but it also means that the relationship between audio and meaning must be learned (Mcgookin, 2020).

A symbolic system is defined as a structured set of signs where the structure may encompass linguistic, visual, musical, or other forms. In such systems, signs possess relationships of contrast and/or combination with each other. This concept aligns with Saussure's (1995) structuralist perspective on semiotics, which emphasizes binary contrasts as fundamental in structural analysis. These binary contrasts, such as nature vs. culture and life vs. death, shape the definition of a concept through its relationship with other concepts within the same system.

Much like the road sign system, as illustrated in Table 3, earcons function as an example of how contrast plays a pivotal role in semiotics. They form a part of a symbolic system and employ auditory cues to convey meaning. In the case of earcons, the auditory elements are manipulated to represent different concepts or actions. For instance, a single beep may indicate one action, while a double beep signifies another. This use of contrasting auditory elements enables users to distinguish between different earcons and understand their associated meanings.

Just as road signs use colors and shapes to create contrast, earcons use variations in sound patterns and rhythms to differentiate between different messages.

So, earcons are similar semiotically to color and shape in road sign system since they are both symbolic in nature. Hence the need to create a system of signs based on contrast.

2.3.3 Linguistics

In the study of linguistics and its underlying semiotics, Ferdinand de Saussure is a central figure, along with Charles Sanders Peirce. Saussure's ideas on linguistics gained prominence in 1916 with the publication of his work titled "*Cours de Linguistique Générale*".

During the nineteenth century, linguists were focused on the study of the transformations languages to explain linguistic changes: linguistics was considered historical or diachronic (Saussure, 2011). Saussure (2011) advocated for a synchronic perspective in the study of linguistics, which emphasizes the analysis of a language at a particular point in time, rather than its historical development. Nonetheless, the author acknowledged the interdependence of both approaches - the synchronic and the diachronic - in understanding linguistics. One noteworthy aspect of the "Saussurean" model is its semiotic nature, which examines language as a tool for constructing meaning. A word encompasses not only a name for a thing but also a combination of a sound image and a concept (Saussure, 2011).

Linguistics studies the main modality of systems, the natural languages, which are the most widely used and highly developed form of communication (Borba, 2003). According to Kracht (2007), "linguistics is sets of signs. Signs combine an exponent (a sequence of letters or sounds) with a meaning. Signs combine a form and a meaning, and they are identical with neither their exponent nor with their meaning." A sign is a pair consisting of a signifier/exponent and a signified/meaning (Saussure, 2011). For example, in English, the string /dog/ is a signifier, and its signified is doghood (Kracht, 2007).

Modern linguistics is based on functional language theory and divides language into the following components (Kracht, 2007; Hickey, 2005):

- Phonetics - studies the set of all human sounds. Focuses on the emission, transmission, and reception of sound - respectively, articulatory, acoustic and auditive phonetics.
- Phonology - concerned with a subset of human sounds and the classification of said sounds. Studies the relationships between phonemes, the smallest abstract unit of sound.
- Morphology - the study of words within a language including their internal structure and relationships to one another. The smallest unit in a language that has meaning is known as a morpheme. Every word is constructed out of one or more morphemes.
- Syntax - rule set of the language under study, concerned with how words and morphemes can be combined into larger expression units, often referred to as sentences.
- Semantics - study of the encoded meaning in a sentence.
- Pragmatics - studies the meaning of information conveyed by a sender and receiver under a specified context, as detailed in the following title.

2.3.3.1 Pragmatics and Sound Design

Pragmatics, as defined by Searle (1969), is the study of language in use and the contexts in which it is used. It involves understanding the choices speakers make in given situations and how listeners interpret these choices. This understanding is crucial in the field of sound design for user interfaces, where sounds serve as a form of language, conveying specific messages to the user. A central concept in pragmatics is the speech act theory (Austin, 1962; Searle, 1969) that posits that when we use language, we perform various kinds of acts simultaneously.

Austin (1962) divided speech acts into locutionary acts (the act of saying something meaningful), illocutionary acts (the intended impact of the statement on the listener), and perlocutionary acts (the actual effect of the statement on the listener).

In the context of sound design, these acts can be understood as follows:

The locutionary act is the sound itself, as produced by the interface. This is the physical sound that the user hears, created by specific combinations of pitch, volume, timbre, and rhythm. The design of this sound is a critical aspect of user interface design, as it must be carefully crafted to convey the intended message effectively.

The illocutionary act is the intended message or function of the sound, such as alerting the user to a new message or warning them of a system error. This is where the designer's intent comes into play. The sound must be designed in such a way that its purpose is clear to the user, whether it's to draw attention to a critical system alert or to provide positive feedback for a completed action.

The perlocutionary act is the user's response to the sound, which could range from performing a specific action to experiencing a certain emotional response. This is the goal of the sound design - to elicit the desired response from the user. This could be as simple as prompting the user to click on a notification, or as complex as evoking a sense of satisfaction or accomplishment.

Searle further categorizes speech acts (or categories of illocutionary act) into five categories: assertive, directive, commissive, expressive, and declarative (1976). These categories provide a framework for understanding and designing the communicative functions of sounds in user interfaces.

For instance, a directive sound in a user interface could be used to instruct the user to perform a certain action, such as saving their work before the system shuts down. An expressive sound, on the other hand, could convey the system's status or reaction to the user's actions, such as a success sound indicating that a task has been completed successfully.

Understanding these pragmatic principles is crucial for effective sound design. By considering the intended illocutionary act and potential perlocutionary effects, designers can create sounds that clearly and effectively communicate with the user.

In the broader context of this research, pragmatics provides a theoretical foundation for understanding how sounds function as communicative acts in user interfaces. This understanding informs the analysis and critique of existing sound designs, as well as the development of new sound design strategies and solutions. It allows us to approach sound design not just as an aesthetic or technical challenge, but as a form of communication, with its own set of rules and principles. This perspective opens new avenues for exploration and innovation in sound design. For instance, by applying categories of illocutionary acts to the design of sounds, we can create more nuanced and effective sound cues. A sound that is designed as a directive, for example, could be more forceful and urgent, prompting the user to take immediate action. On the other hand, a sound designed as an expressive could be more subtle and nuanced, conveying a range of possible system states or responses.

Furthermore, by considering the perlocutionary effects of sounds, we can design sounds that not only communicate effectively but also elicit the desired responses from users. This could involve designing sounds that motivate users to engage more deeply with the interface or sounds that create a more positive and satisfying user experience.

The application of pragmatics to sound design also has implications for the evaluation and critique of sound designs. By providing a clear framework for understanding how sounds function as communicative acts, it allows us to evaluate sounds not just based on their aesthetic or technical qualities, but also based on their effectiveness as forms of communication. This could lead to more rigorous and comprehensive evaluations of sound designs, and ultimately to the development of better sound design practices and standards.

To summarize, several interconnected fields collaborate to shape how information is conveyed through sound. Auditory displays and sonification, for instance, utilize sound to communicate data relationships, encompassing principles from psychology, computer science, and audio engineering. These disciplines merge to create compelling auditory experiences, highlighting the convergence of diverse knowledge.

Auditory cues, analogous to visual icons, transmit information through sound. These cues are crucial components of auditory displays, where their design draws from principles of perception, acoustics, and psychoacoustics. The field of psychoacoustics, dedicated to understanding how we perceive sound, provides insights into both objective and subjective aspects of sound perception. This knowledge guides the design of auditory cues and displays, ensuring they are effective and emotionally resonant.

Sounds carry information about the world. When listening to sounds (speech, music or product sounds), communication takes place. Each acoustic event can be perceived as a sign carrier through which information is communicated. So, sound designers are, in a sense, engineers

of communication (Jekosch, 2015, p.193). Semiotics, the study of signs and their meanings, enriches the narrative by dissecting the relationships between auditory cues and their significance. In this research, we propose that linguistic principles may contribute to a structured approach to sound design. The language components - phonetics, phonology, morphology, syntax, semantics and pragmatics - can frame a heuristic to describe, evaluate, create or improve systems of sound signals for devices. Pragmatics (studies language in use and the contexts in which it is used) categorizes illocutionary acts into five categories: representatives, directives, commissive, expressive, and declarative. We suggest that they may provide a reference framework for organizing, understanding, and designing the communicative functions of sounds.

The intersection of these fields is profound. Auditory displays and sonification benefit from the insights of psychoacoustics, ensuring that the auditory messages are accurately conveyed and perceived. Auditory cues, shaped by semiotics, are designed to evoke specific meanings, resonating with users on a cognitive and emotional level.

Ultimately, this intricate web of disciplines transforms auditory communication from simple data transmission into a multi-dimensional experience. By combining elements of sound, perception, meaning, and emotion, these fields collaboratively craft auditory interactions that are informative, captivating, and thoughtfully designed.

3. Methodology

This section aims to outline the methodical creation of abstract sounds that convey meaning and are easy to learn while maintaining an aesthetic consistency. We will delve into the principles and guidelines that underpin this methodology. We will explore how to approach sound design systematically, addressing key questions about the types of messages we can create, the responses they can trigger, and the factors that facilitate or hinder the assimilation of sound signals.

Our goal is to provide a heuristic that allows for a better description, evaluation, production, and optimization of sound signals, enhancing their effectiveness and accessibility.

This section is composed of three (3) chapters: guidelines, pragmatics for a hypothetical game and sound design model. In the guidelines chapter we investigate how to connect linguistics and sound design and create guidelines for conveying messages with abstract sounds. In the pragmatics for a hypothetical game chapter, we delve into how we can use Searle's speech act theory to classify the different messages created by abstract sound systems. In the sound design model chapter, we look into how we can organize information.

Like road sign systems that utilize color, shape, and symbols to convey crucial information to drivers, this methodology harnesses the power of abstract sounds as a mode of communication. Just as road signs are designed to be universally understood regardless of language or cultural background, abstract sounds offer a similar potential for cross-cultural and cross-linguistic communication.

The methodology draws insights from various disciplines to enrich its foundations. Psychoacoustics, the study of how humans perceive sound, and sonification, the conversion of data into sound, play vital roles. These disciplines provide reference points from the realm of sound and music, offering best practices for creating sounds that are pleasant, effective, and evocative. This knowledge allows sound designers to make informed decisions about what acoustic parameters could be best for each given situation.

Furthermore, semiotics, the study of signs and symbols, and linguistics, the study of language, provide valuable guidelines for crafting meaningful and comprehensible abstract

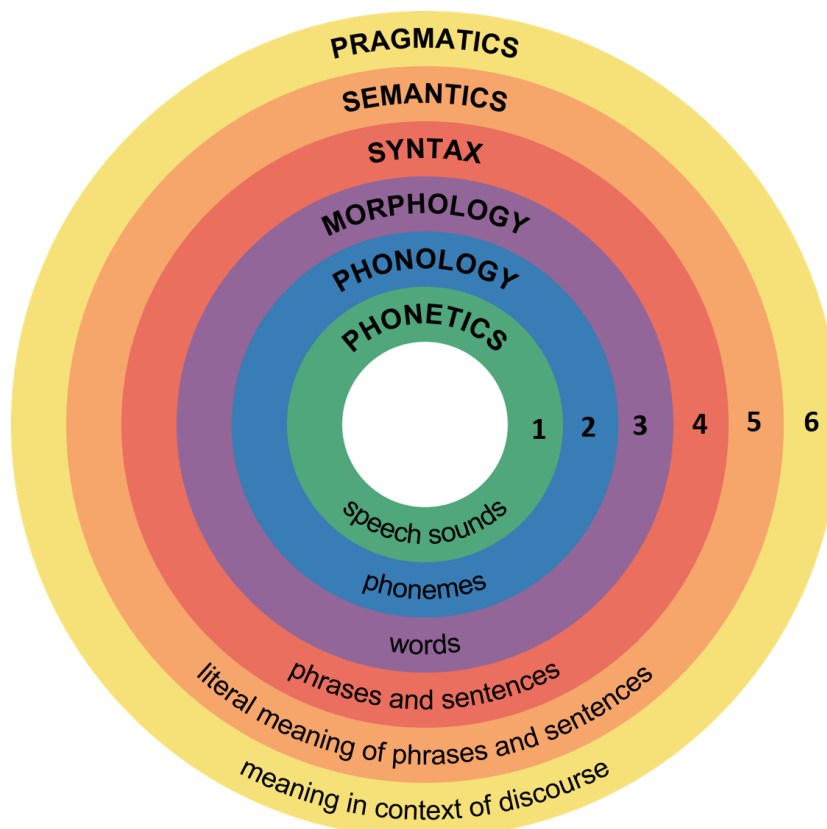
sounds. Just as semiotics helps designers understand how visual symbols convey meaning, it can also inform the creation of auditory symbols (earcons) using abstract sounds. Linguistic principles offer a structured approach to sound design, enabling designers to imbue sounds with specific meanings and emotions, akin to the way words are used in language.

3.1 Guidelines

In this chapter, we connect linguistics and sound design by developing guidelines for conveying messages with abstract sounds. The goal is to contribute for a heuristic to describe and evaluate the production and optimization of systems of messages transmitted through sound, based on the parallels between semiotics, linguistics, and sound design, interpreting how sounds can act as signs that represent something else.

The proposed guidelines follow the structure of linguistics as a model to evaluate and develop non-speech sounds that, just like language, systematically convey meaning to the user. Figure 7 (image is in the public domain, adapted) presents six levels of linguistic analysis that range in depth between the specifics of the sounds we make to form language to the context surrounding speech events: phonetics, phonology, morphology, syntax, semantics, and pragmatics (from most specific to the broadest).

Figure 7 - Major levels of linguistic structure



We can thus propose six groups of fundamental questions to consider:

1. **Phonetics** - What are the material conditions in the production of sound? Which parameters describe it and can be shaped for its production? What are the conditions of production imposed by the transmitting apparatus, and what are the conditions of reception inherent to the user? These questions are dealt with by phonetics, acoustics and psychoacoustics.

In the musical domain, phonetics would refer to all the sounds a specific system could potentially reproduce regardless of if they are used by the system to convey meaning or not.

2. **Phonology** - What properties, and in each of them, what range of possible attributes, does the system select for the creation of differentiated units of meaning? What are the minimal material units that can be used to produce differentiated senses (even though they themselves do not yet hold meaning)? That is, what sounds, and acoustic properties of sounds the system selects to establish meanings? This is the question that differentiates phonology from phonetics, and that will deal with phonemes. It is also here that lies the questions of how the user can perceive physical differences, in what situations and under what conditions.

In the context of abstract sounds like earcons, differentiated units of meaning are established through variations in sonic elements. These variations convey distinct messages or actions in user interfaces.

3. **Morphology** - What is the lexicon of the system? What are the morphemes, composed of phonemes, that the system gathers in its "dictionary" and that can be messages in themselves or serve to compose other messages? This is the problem addressed by morphology. It is also here that one can begin to address the question about the relationship that is established between the signifier and the thing referred to (of similarity in auditory icon, or of symbol convention (earcon, or even of causal inference from an index) and how this can facilitate or hinder the learning curve.

Earcon morphology addresses how these individual sounds combine to create meaningful earcons, resembling the composition of morphemes in language.

4. **Syntax** - What are the types of morphemes and what are the legitimate combination rules for producing new messages by combining morphemes? These are questions addressed by syntax and by Chomsky's generative grammar (1965, 1995), which allows us to group morphemes used to express events (such as receive, emit, create, delete, submit, download, connect, disconnect, change, open, close, copy, etc.), the qualification of events (success, failure, accomplish, in process, urgent, etc., as well as locations (as in email, messenger, alarm, etc.), subjects (user, device, particular signs to indicate people, etc.) and to qualify them. We can recognize here verbal, adverbial as well as noun and adjectival phrases.

Applying syntax to earcons means defining rules that govern the permissible combinations of sound elements, ensuring coherent and meaningful earcon formation.

5. **Semantics** - Having delimited the morphemes and the syntax, what are the messages that can be generated by the system and interpreted by the users? Are the messages closed or can they be open, being created by potentially unlimited combinations? What messages are possible and what messages are necessary? Does this step come at the end, or should it be at the beginning delimiting the Sound Designer's field of action? These are the questions addressed by semantics (Eco, 1999). It is also here that the problem of the user's acquisition of mastery over the code can be posed in its full extent. How can the user internalize morphemes and syntactic rules to understand the semantics in each message? Can this be the result of a passive association, by habit, or is it the result of active learning, through education? What factors facilitate and hinder this assimilation, from the importation of already existing conventions to the use of similarities in sounds?

Earcon semantics involves the interpretation of meaning behind specific sound sequences. Users attribute certain actions or interface responses to particular earcons based on learned associations.

6. **Pragmatics** - Finally, what kinds of messages can we have and what kind of response can they trigger? Are they assertive messages, carrying only information, or directive messages, requiring an action? Does it make sense to take into consideration expressive and declarative or even commitment messages? These are questions dealt with by Pragmatics, first formalized by Austin (1962) and Searle (1969, 1976). Searle set up the following classification of illocutionary speech acts:

- Assertives - commit a speaker to the truth of the expressed proposition. Psychological state: belief.
- Directives - cause the hearer to take a particular action, e.g. requests, commands and advice. Psychological state: want, wish, desire.
- Commissives - commit a speaker to some future action, e.g. promises and oaths. Psychological state: intention.
- Expressives - express on the speaker's attitudes and emotions towards the proposition, e.g. congratulations, excuses and thanks. Psychological state: range of different possible states.
- Declarations - change the reality in accord with the proposition of the declaration, e.g. baptisms, pronouncing someone guilty or pronouncing someone husband and wife. Psychological state: none.

These types of messages may present modes (ok, not ok, neutral) and modes, in turn, may present moments (before, now, in the future). These six groups of questions seem to provide a systematic approach to the creation of a heuristic that allows a better description and evaluation,

as well as the creation and optimization of systems of sound signals to be incorporated in devices, which can be used in isolation, as a complementary way to visual signs, or substitutes, addressing from the outset the important issue of accessibility for people with limitations of other senses such as sight or even touch.

For the future there needs to be further work done to contextualize the existing literature into this new system as well as further analyses of the field of linguistics and semiotics.

3.2 Pragmatics for a hypothetical game

This chapter expands the role of pragmatics and speech acts on the methodology.

Table 4 represents an application of speech act theory. The table is filled up with sound cues for a hypothetical game. It categorizes different sound cues according to their illocutionary force (assertive and directive) and their state in the game (in, in process, and out). It also identifies the source of the sound cues (singular, plural, and neutral). For this methodology, we will focus on assertive and directive speech acts as these are the most used in auditory displays.

Table 4 - Speech act table for hypothetical game

	Assertive			Directive		
	Ok	Nok	Neutral	Ok	Nok	Neutral
In	Picks up positive item	Picks up negative item	Picks up neutral items	Tutorial: do to win	Will loose recording	Put name
In process	Starts to be able to eat	Losing life	Walking	Must collect coins against time	Must run or die	Must choose room
Out	Positive hit on thing	Negative hit on thing	No hit	Order to do	Order not to do	Suggestion
	Singular	Plural	Neutral	Singular	Plural	Neutral
1st	Player shoots	Player's team shoots	Familiar environment	Player orders	Player forbids	Player suggests
2nd	Enemy shoots	Enemy's team shoots	Enemy environment	Enemy orders	Enemy forbids	Enemy suggests
3rd	Neutral entity	Group of neutral entities	Neutral environment	Someone informs	Someone warns	Someone suggests

In the assertive category, sound cues represent a situation in the game. For instance, the sound of picking up a positive item signifies a beneficial event (ok), while the sound of picking up a negative item signifies a detrimental event (nok). Neutral sounds, such as the sound of picking up neutral items, do not signify a particularly beneficial or detrimental event.

In the directive category, sound cues are intended to get the player to do something. For example, a sound cue during a tutorial instructs the player to perform a certain action to win (ok), while another sound cue warns the players that they will lose their recording if they do not act (nok). Neutral directives, such as the sound cue to put a name, do not urge immediate action or carry a sense of urgency.

The 'in', 'in process', and 'out' categories represent different states or stages in the game. 'In' sounds are associated with the initiation of an action or event, 'in process' sounds are associated with an ongoing action or event, and 'out' sounds are associated with the conclusion or outcome of an action or event.

The 'singular', 'plural', and 'n' categories represent different sources of sound cues. 'Singular' sounds are produced by individual entities (e.g., the player or an enemy). 'Plural' sounds are produced by groups of entities (e.g., the player's team or the enemy's team). 'Neutral' sounds are produced by the environment or non-specific sources.

This table provides a framework for designing and interpreting sound cues in a game environment, based on Searle's speech act theory. It shows how the illocutionary force, state, and source of a sound cue can be used to convey specific messages and elicit specific responses from the player. This approach to sound design can contribute to a more immersive and intuitive experience and aligns with the overall aim of this thesis to explore the role of sound in communication.

The application of Searle's speech act theory to game sound design also opens new avenues for research and innovation. For instance, future work could explore how different combinations of illocutionary force, state, and source can be used to create more nuanced and effective sound cues. It could also investigate how players interpret and respond to these sound cues, and how these responses influence their overall gaming experience.

3.3 Sound Design model table

In this chapter, a model for sound design table is presented. The goal is to organize the data we have about the methodology in a systematic and structured manner, helping the sound designer to develop a visual framework of all the various components that go into analyzing and creating sounds that convey meaning.

We established a parallel between the road sign system and the auditory field, which resulted in Table 5. It is the “template” for a table that can be expanded or adapted according to the specifics of each situation.

Table 5 - Model for User Interface Sound Design: 6 types and characteristics

Type of Sound	Pitch	Timbre	Rhythm	Duration	Envelope	Sign-Perception
Warning Sounds	High	Sharp	Staccato	Short	Fast	To warn and inform the user of a danger
Error Sounds	High	Sharp	Staccato	Short	Fast	To inform the user of an error or a problem
Confirmation Sounds	Medium	Neutral	Steady	Short	Slow	To confirm a user's action or a system's response
Feedback Sounds	Medium	Neutral	Varying	Short	Slow	To provide feedback on a user's action or a system's response
Regulatory Sounds	Medium	Neutral	Steady	Medium	Slow	To convey obligations, restrictions or prohibitions
Informative Sounds	Low	Smooth	Legato	Long	Slow	To provide travel information (direction, distance, place)

It is built on the guidelines of the methodology being developed. It deals with one-element - earcons - designed for the purpose of general auditory displays.

The 'type of sound' column is used to categorize the different types of sounds based on their purpose or meaning. The columns 'pitch', 'timbre', 'rhythm', and 'duration' are used to describe the characteristics of sound, like the way street signs use shape and color to convey meaning. Their range of values and associated adjectives are based on general principles of psychoacoustics that will be used to give concrete values to the adjectives. The types of sound characteristics and their ranges of values were chosen taking into consideration the studies of psychoacoustics and the guidelines for developing earcons (Brewster et al., 1995a; Patterson, 1982; Kuwano et al., 2000) as well as the concept of contrast given to us by Saussure (1995).

The parameters chosen for the figure considered the parameters that the reviewed literature found best conveyed meaning to users. So, they could correctly tell when there was a change in the parameter and their perception changed accordingly. Parameters such as loudness were not considered as, although the users could easily identify a difference and act accordingly, there is usually the possibility to change volume, as noted in the auditory cue design guidelines (Brewster et al., 1995a).

The values chosen for the warning signal consider the studies of warning signals as these showed a clear range of values that successfully conveyed a sense of warning cross culturally as well as efficiently minimizing the possibility of the masking effect. The type of values for other types of sounds were mostly developed in contrast to the values associated with warning signals

following the concept. The 'signal perception' column is used to convey the objective or intended purpose of the sound, like the way street signs convey information and directions to road users.

As mentioned before, this was only an example of a possible table that could adapt to different situations. It is worth noting that there is a possibility for greater detail, as demonstrated in Table 6. In this case, we added the characteristics 'spatialization' and 'volume'.

Table 6 - Model for User Interface Sound Design: 6 types and 8 characteristics

Type of Sound	Pitch	Timbre	Rhythm	Duration	Spatialization	Volume	Envelope	Sign-Perception
Warning Sounds	High	Sharp	Staccato	Short	Directional	Loud	Fast	To warn and inform the user of a danger
Error Sounds	High	Sharp	Staccato	Short	Directional	Loud	Fast	To inform the user of an error or a problem
Confirmation Sounds	Medium	Neutral	Steady	Short	Ambient	Moderate	Slow	To confirm a user's action or a system's response
Feedback Sounds	Medium	Neutral	Varying	Short	Ambient	Moderate	Slow	To provide feedback on a user's action or a system's response
Regulatory Sounds	Medium	Neutral	Steady	Medium	Ambient	Moderate	Slow	To convey obligations, restrictions or prohibitions
Informative Sounds	Low	Smooth	Legato	Long	Spatial	Soft	Slow	To provide travel information (direction, distance, place)

Table 7 is a simpler version.

Table 7 - Model for User Interface Sound Design: 3 types and 4 characteristics

Type of Sound	Pitch	Timbre	Duration	Sign-Perception
Danger Warning Sounds	High	Sharp	Short	To warn the user of a danger or potential hazards
Regulatory Sounds	Medium	Neutral	<u>Medium</u>	To convey obligations, restrictions or prohibitions
Informative Sounds	Low	Smooth	Long	To provide information or guidance to the user

This kind of table might not always be necessary or even useful in the creation of sound. We will see later other ways to approach sound design. However, if you want to make a system of abstract sounds that convey meaning with a systematic approach or analyze other similar systems it is a valuable tool. It allows the categorization of sounds based on their purpose or meaning, and the delineation of their characteristics such as pitch, timbre, rhythm, and duration.

However, it is crucial to note that the applicability and utility of this table are contingent on the specific approach taken towards sound design. As will be demonstrated in the chapter dedicated to practical cases, not all sound design scenarios align with the systematic categorization suggested by the table. For instance, the creation of transformative earcons, a form of data sonification, involves a more intricate process that associates different musical parameters to each message type.

4. Practical cases

The easiest way to put the methodology into testing seemed to be by analyzing existing cases. Not all the approaches developed in the methodology need to be applied to every situation. Sometimes, it will not give us much information on a particular case, or we do not have enough data to make an assessment. The practical cases constitute of three different case types, each one meant to give a different insight into the methodology and help better understand its potential in possible use cases: simple sound systems, smartphones and transformative earcons. Audio examples may be found in annex “[Practical cases sounds](#)”.

4.1 Simple sound systems

In this chapter, an in-depth analysis of simple systems that utilize audio to communicate messages will be conducted, with a particular focus on the cardiac arrest machine. Additional examples of straightforward auditory systems will also be presented. However, unlike the cardiac arrest machine, these examples will not be subjected to a detailed analysis based on linguistic layers or the speech act classification. This is due to their significant similarities with the cardiac arrest machine. Instead, a brief description of each case study will be provided, accompanied by a graphical representation for easy reference.

The importance of presenting examples of simple, everyday sound systems that relay information cannot be overstated. They help to demonstrate the practical applications and significance of auditory systems in daily life. By highlighting these instances, the pivotal role that auditory systems play in communication, information transfer, and overall human interaction is underscored.

These examples not only emphasize the relevance and importance of understanding and studying auditory systems in specialized fields, but they also state their significance in broader contexts. Thus, the study of these auditory systems is crucial to understanding how, every day, communication and interaction occur. We will apply the guidelines based on linguistic layers and

pragmatics speech acts to the following simple sound systems: cardiac arrest machine, car reverse gear, elevator, and traffic light crossing (audio examples may be heard in annex “Practical cases sounds: 1 - Simple sound systems”).

4.1.1 Cardiac Arrest Machine

4.1.1.1 Guidelines based on linguistic layers

1. **Phonetics** - Involves the actual sounds produced by the cardiac arrest machine. These are typically beeping sounds, with the material conditions in sound production involving the hardware and software of the machine that generate these sounds.

2. **Phonology** - The phonology of the cardiac arrest machine is defined by silence and sound, and duration that combine to form rhythm.

3. **Morphology** - Concerns the meaning created by the sounds chosen in phonology. In the cardiac arrest machine, a medium rhythm indicates that the machine is working, and the heartbeat is regular. A fast rhythm, on the other hand, indicates an abnormal heartbeat.

4. **Syntax** - There is no information concerning syntax since there is no combination of messages or different message categories.

5. **Semantics** - In the case of the heart rate monitor, there are no complex messages being displayed. Therefore, there are no elements to analyze concerning this layer. The machine can only give us one message at a time. There is no combination of messages.

6. **Pragmatics** - In the context of the cardiac arrest machine, the illocutionary acts are primarily assertive, meaning they declare a state of affairs. The sound cue associated with a regular heartbeat is an in-process positive assertive act that communicates the message that the patient's heart is functioning normally. The sound cue for an abnormal heartbeat is a negative assertive act in process that communicates an irregularity or potential issue with the heart's rhythm.

The sound cue for a cardiac arrest, represented by a sustained note, is also a negative assertive act. It communicates a critical situation, namely that the patient is experiencing a cardiac arrest. This sound cue is designed to have the perlocutionary effect of triggering immediate action from the medical staff.

In this context, the speech acts performed by the cardiac arrest machine are not complex or varied. They are designed to convey specific, critical information about the patient's condition in a clear and unambiguous manner. The effectiveness of these speech acts lies in their simplicity

and clarity. They do not require interpretation or inference on the part of the medical staff. Instead, they provide direct, actionable information that can guide the medical staff's response.

4.1.1.2 Cardiac arrest table

Table 8 shows how different sound cues are used to represent different states of the cardiac arrest machine. The “message” column represents the meaning created by the sounds, the “rhythm” column represents the properties selected by the system to create differentiated units of meaning, the “purpose” column represents the intended response from the listener, and the 'icon type' column represents the type of auditory icon used. The “speech act” column categorizes the sound cue according to its illocutionary force, state, and mode.

Table 8 - Cardiac arrest

Message	Rhythm	Purpose	Icon type	Speech act
Regular heartbeat	Medium	Indicate that the machine is working, and the heartbeat is stable	One-element earcon	Assertive In process ok
Abnormal heartbeat value	Fast	Indicate abnormal heartbeat	One-element earcon	Assertive In process nok
Cardiac arrest	Sustained note	Indicate cardiac arrest	One-element earcon	Assertive In nok

In the context of the present research, this table applies various linguistic categories to the design of sound cues in a cardiac arrest machine. It shows how different types of sounds can be used to convey specific messages and elicit specific responses, contributing to overall effectiveness.

4.1.2 Car Reverse Gear

The reverse gear in a car typically produces a high-pitched beeping sound. This sound is continuous while the car is in reverse, indicating that the car is moving backward. When the car gets close to an object, the rhythm of the sound increases to a fast pace, indicating that the car's sensor has detected an object. If the car hits an object, a sustained note is produced, indicating that the car has hit an object. These sounds serve as auditory cues, providing important information about the car's movements and surroundings. Table 9 analyses the sound messages of cars reverse gear.

Table 9 - Reverse gear

Message	Rhythm	Purpose	Icon type	Speech act
Engaged reverse gear	Medium	Indicate that the car is reversing	One-element earcon	Assertive In process ok

Close to an object	Fast	Indicate the car sensor has detected an object	One-element earcon	Assertive In process nok
Very close to an object	Very fast	Indicate the car sensor has detected an object very close	One-element earcon	Assertive In process nok

4.1.3 Elevator

Elevators often produce a variety of sounds to indicate their status, as shown in Table 10. When the elevator arrives at a floor, a single beep is produced. Elevators may produce one beep for going up and two beeps for going down. These sounds serve as simple earcons, providing important information about the elevator's status and movements.

Table 10 - Elevator status

Message	Rhythm	Purpose	Icon Type	Speech Act
Arrived at floor (going up)	Single beep	Indicate that the elevator has arrived at a floor and is going up	One-element earcon	Assertive In process ok
Arrived at floor (going down)	Double beep	Indicate that the elevator has arrived at a floor and is going down	One-element earcon	Assertive In process ok

4.1.4 Traffic Light Crossing (Australian Crossings)

Australian traffic light crossings produce a variety of sounds to indicate their status as presented in Table 11. When it is safe to cross the road, a rapid ticking sound is produced. This sound is so iconic that it was sampled by Billie Eilish in her Grammy award-winning song, "Bad Guy". When it is not safe to cross the road, a continuous beep is produced. When pedestrians must wait to cross, there is silence. These sounds serve as simple earcons (one element), providing important information about when it is safe to cross the road.

Table 11 - Traffic light crossing

Message	Rhythm	Purpose	Icon Type	Speech Act
Safe to cross	slow	Indicate that it is safe to cross the road	One-element earcon	Directive in process ok
Don't walk	fast	Indicate that the pedestrian must wait to cross	One-element earcon	Directive in process nok

The previous examples illustrate how illocutionary force, state, and source of simple sound systems can be used to convey specific messages. All communicated their purpose (a specific meaning or function) through one element earcon - abstract, synthetic tones that create structured auditory messages with a single mapping between one motive and the information it represents -

establishing a direct and unambiguous mapping between motive and meaning. This unique characteristic simplifies the communication process, leaving little room for interpretation or misinterpretation.

The illocutionary forces present (the intended impact of the statement on the listener) were mainly assertive and directive. Applying the methodology's guidelines for linguistic analysis to simple sound systems design may be useful to identify and/or craft earcons that assertively declare the intended message. These messages are precisely categorized into 'in,' 'in process,' and 'out' states, enabling users to grasp timing and progression effortlessly. Additionally, the differentiation of sources as 'singular,' 'plural,' or 'neutral' aids users in attributing sounds to specific entities or environmental factors within the system.

To sum up, simple sound systems' effectiveness lies in the use of “minimalist” elements such as one element earcon and the adequate illocutionary force to provide clear, critical, direct, trustable and/or actionable information.

Annex “Practical cases sounds: 1 - Simple sound systems” has audio examples of cardiac arrest machine, car reverse gear, elevator, and traffic light crossing earcons.

4.2 Smartphones

In this chapter, we delve into the sound profiles of various smartphone models, including the Vodafone p600, Samsung Galaxy S5, Xiaomi mi a3, and Huawei (Table 12). Audio examples may be heard in annex “Practical cases sounds: 2 - Smartphone”.

Table 12 - Smartphones sounds

	Dock	Undock	Lock	Unlock	Keypresses *5
Vodafone p600	ascending C G C arpeggio with a sine type timbre (longer release)	descending C G C arpeggio with a sine type timbre	water drop	same water drop	Typewriter sound with no invalid sound
Samsung/Galaxy-S5 system	ascending C G C arpeggio with a sine type timbre (sharper)	descending C G C arpeggio with a sine type timbre	2 water drop or water drop and splash	1 long water drop	Typewriter sound except for invalid Bb sine
Xiaomi mi a 3	Low pitched tapping sound	N/A	medium range tapping sound	N/A	Clicking tapping sound except invalid Bb sine
Huawei P10	Low pitched tapping sound (same as Xiaomi mi a3)	error sound sine C	Water sound tuned low Eb	water sound midrange (1000 to 3500)	Typewriter sound except for invalid Bb sine

Each model was scrutinized based on five sound parameters: 'dock', 'undock', 'lock', 'unlock', and 'keypresses'. However, our initial approach of employing a grid to analyze these sounds was found to be less efficient due to the lack of consistency in the attribution of acoustic parameters to these actions. Hence a table simply describing each sound was made instead.

The inconsistencies posed a challenge in recognizing the meaning of each sound and differentiating between phone models based on their sound profiles. To better understand the complexity of the sound design, we turned to linguistic analysis and broke down the sound design into several layers:

1. **Phonetics** - This layer pertains to the physical properties of the sounds that the smartphones can produce. In general, smartphones start to roll off from 1k below with the frequency response with the slightly varying from model to model.
2. **Phonology** - This layer encompasses the process of selecting and organizing sounds to effectively convey meaning. In the context of smartphones, distinct sounds are employed to indicate various actions. However, a lack of consistency was observed in assigning specific sounds to each phone model, as there was no identifiable acoustic parameter, sound characteristic, or family associated with them. For instance, when it comes to the docking and undocking actions in both the Vodafone and Samsung models, an ascending arpeggio is consistently used to signify docking, while a descending arpeggio with a distinct timbre indicates undocking. This consistent approach effectively conveys the meaning of these two contrasting actions. However, the lock and unlock actions deviate from this cohesive system. Both models use water drop sounds, which undermines the establishment of a clear contrast and hampers any distinctive sound identity the models or brands could have. Moreover, only the Samsung model features a different water drop sound, further adding to the inconsistency. Furthermore, there are instances where certain sounds are shared across multiple models, giving the impression that the sounds were randomly selected from a common sample pack. This diminishes the uniqueness and individuality of each model's sound design. In contrast to the inconsistencies, the Xiaomi model stands out by consistently utilizing tapping sounds for all actions. However, the challenge lies in distinguishing between these tapping sounds and associating each one with a specific action. This lack of clarity disregards the idea of contrast, which is vital for effective communication of meaning. It is akin to having road signs all rendered in varying shades of blue instead of utilizing highly contrasting colors, thereby impeding quick and accurate comprehension.
3. **Morphology** - This layer deals with the meaning created with the parameters selected by phonology. As stated before, an ascending arpeggio in Vodafone and Samsung models represents the action of docking, while a descending arpeggio signifies undocking. Water drop sounds are used to indicate both lock and unlock actions. In the Xiaomi model, all actions are represented by different tapping sounds. All models apart from Xiaomi use

typewriter sounds for keypresses. Also, all of them used a Bb tuned sine wave to represent error sound apart from Vodafone which has no invalid sound.

4. Syntax - N/A
5. Semantics - N/A
6. **Pragmatics** - This layer concerns the use of language in social contexts and the effects of language on perception and behavior. In the case of smartphones, it refers to the intended effects of the sounds on the user. For example, an ascending arpeggio, which is usually associated with positive messages, is used for docking and unlocking (assertive out ok), while a descending arpeggio, which is often associated with negative messages, is used for undocking, and locking (assertive out not ok).

The sound design of these smartphones could greatly benefit from a more systematic approach. By establishing a consistent theme or pattern for the sounds, it would be easier to attribute specific meanings to the sounds. This would not only enhance the user experience but also improve the aesthetic consistency of the sound design.

In the following chapter, we will analyze another sound design model for smartphones. This model will aim to establish a consistent theme or pattern for the sounds, thereby making it easier to attribute specific meanings to the sounds enhancing the overall user experience.

As explained, this research is focused on systematic creation of abstract sounds that convey meaning, are easy to learn, and maintain aesthetic consistency. This is achieved through the application of linguistic analysis and the principles of psychoacoustics to sound design. The sound design model table, which categorizes and describes the characteristics of sound, is a key tool in this methodology.

The application of Searle's speech act theory to sound design provides a new perspective on the creation and interpretation of sound cues. By considering the illocutionary force, state, and source of a sound cue, we can create more nuanced and effective auditory signals. This approach not only enhances the user experience but also contributes to the broader goal of advancing in our understanding of sound in user interfaces.

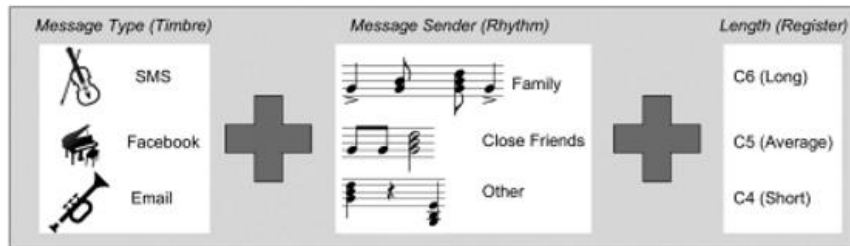
The analysis of smartphone sound profiles revealed a lack of consistency in the attribution of acoustic parameters to actions. This inconsistency makes it difficult for users to recognize the meaning of each sound and differentiate between phone models based on their sound profiles. However, by adopting a systematic approach to sound design, we can maintain aesthetic consistency, improve the user experience, and make it easier to attribute specific meanings to the sounds.

In the next chapter, we will analyze a hypothetical sound design model for smartphones that uses transformative earcons as a basis and better accomplishes the goal of conveying meaning through with a systematic approach.

4.3 Transformational earcons

In Figure 8 (McGookin, 2019), we see the so called “transformational earcons”, a form (a more complete example) of data sonification.

Figure 8 - Foundations in SD for Embedded Media: a Multidisciplinary Approach



The image is a diagram that illustrates the process of creating transformative earcons. It shows three categories: 'message type', 'message sender' and 'length'. Each category has different options, and by combining these options, a set of 27 earcons can be generated. Audio examples were developed based on the diagram. These can be heard in annex “Practical cases sounds: 3 - Transformational earcons”.

The transformative earcons are a type of auditory cue, which are abstract, synthetic tones that can be used to create structured auditory messages. They are designed based on musical principles and employ musical motives as their core form.

In the figure, the transformative earcons are being used to represent different types of messages. The type of message, its sender, and its length are the parameters that are being transformed into sound. For this example, we associate a musical parameter to each message type.

- Timbre - Different timbres are used to represent different message types. This helps to prevent confusion and contributes to more effective earcons.
- Rhythm and Pitch - The rhythm and pitch of the earcons are used to represent who the message is from (the sender).
- Register - The register (tessitura) of the earcons is used to represent the length. Long messages have a high register, while short messages have a low register.

This is a good example of how a systematic approach can create more intricate messages, streamline the creation process, and give consistency to the sounds.

In this case, the linguistic layer “semantics” applies, since multiple units of meaning are combined to form more complex messages. Through the sound, we can communicate that we got an SMS from a close friend that was short and other combinations. This approach not only allows communication of more complex information to the user but also for a more streamlined process and design, that will retain aesthetic consistency and efficiency.

If we analyze this example according to the linguistic layers:

1. Phonetics - It will depend on the reproduction system of the phone model that the system will be running on. But for most phones, mid to low range mono sounds are usually ideal for reproduction, as the speaker can't faithfully reproduce low sounds or stereo.

2. Phonology - This is about the selection of specific attributes of sound to create differentiated units of meaning. In this case, it is timbre, rhythm, and register, each one with three (3) possible arguments, enabling nine (9) different phonemes.

3. Morphology - The minimal viable unit of meaning that exists. In this case, we know that different timbres mean different message types, different rhythms mean different senders and different registers mean different message lengths, giving us a total of nine (9) morphemes.

4. Syntax - The syntax is the possible types of morphemes and the rules of their combination. So, in this case, the syntax is: message type, message sender and length and how the nine (9) morphemes can combine with each other.

5. Semantics - The meanings of the earcons. Each combination of timbre, rhythm, and register is intended to communicate a specific type of message from a specific sender of a specific length to the user that combines to form a total of 27 different possible messages.

6. Pragmatics - Concerns the intended effects of the earcons on the user. The earcons are primarily assertive neutral, providing information about the type, sender, and length of a message. The expected response is for the user to understand the information conveyed by the earcon and act accordingly, such as by checking a long SMS from a family member.

This analysis provides a systematic approach to understanding the creation and interpretation of transformative earcons. It highlights the importance of each linguistic layer in the design and use of these earcons and provides a framework for further research and development in this area.

Note: audio examples of simple sound systems, smartphones and transformative earcons may be found in annex "[Practical cases sounds](#)".

5. Final considerations

The field of sound design has seen significant advancements in recent years. However, a noticeable gap persists when it comes to systematic approaches in the analysis and creation of sounds that convey information. This study undertook a comprehensive literature review in the fields of auditory representation, psychoacoustics, semiotics, and linguistics to provide a foundation for existing theory. Based on this research, it was proposed a methodology to create and evaluate sounds to communicate. Its effectiveness and reliability were evaluated through practical cases, providing systematic insights into the impact of sounds.

5.1 Objectives satisfaction

By applying some of the linguistics principles (such as language components and illocutionary acts) to earcons (sound) design, this research pursued the following question: *how can we develop a methodology to create and evaluate sounds that communicate effectively?*

Sounds carry information about the world. When listening to sounds, communication takes place. Each acoustic event can be perceived as a sign carrier through which information is communicated. Therefore, sound designers are also communication engineers. Grounded on semiotics, the study of systems of signs based on codes, and linguistics, the study of language, we proposed guidelines for abstract sounds design. Just as semiotics helps designers understand how visual symbols convey meaning, it also contributes to comprehend how auditory symbols (earcons) can act as signs that represent something else. As road signs use colors and shapes to create contrast, earcons use variations in sound properties to differentiate between messages.

Linguistic principles can contribute to a structured approach to sound design. The language components - phonetics, phonology, morphology, syntax, semantics, and pragmatics - framed a

heuristic to describe, evaluate, create or improve systems of sound signals for devices. Pragmatics (studies language in use and the contexts in which it is used) illocutionary acts (assertive, directives, commissive, expressive, and declarative) also provided a reference framework for organizing, understanding, and designing the communicative functions of sounds. By establishing a parallel between the road sign system and the auditory field we also proposed a "template table" for user interface sound design.

In the practical cases, the methodology was applied to simple sound systems (cardiac arrest machine, car reverse gear, elevator, and traffic light crossing), smartphones and transformative earcons. Linguist layers contributed to organize and understand the referred sound systems.

We concluded that illocutionary force, state, and source of simple sound systems can be used to convey specific messages through one element earcon, establishing a direct and unambiguous mapping between motive and meaning. This unique characteristic simplifies the communication process, leaving little room for interpretation or misinterpretation.

Delving into the sound profiles of various smartphone models, we noticed that the sound design of these devices could greatly benefit from a more systematic approach. By establishing a consistent theme or pattern for the sounds, it would be easier to attribute specific meanings to the sounds.

Finally, we looked into an example of transformational earcons that combined three message types with three message senders with three lengths, generating 27 possible earcons, a good example of how a systematic approach can create more intricate messages, streamline the creation process, and give consistency to the sounds. This study aimed to improve experiences across these domains by filling a gap in systematic methodologies for sound analysis and creation.

The systematic approach of this study can enable sound designers to make informed decisions regarding sound selection, synthesis, integration, and evaluation. This way, it will be possible to produce sounds that improve usability, engagement, and user satisfaction. It aims to contribute to enhancing interactions and creating more pleasurable and immersive experiences in a variety of contexts.

5.2 Future work

The current research has certain limitations that need to be acknowledged. The methodology developed is largely based on theoretical principles and needs further empirical testing to validate its effectiveness across various scenarios.

Moreover, sound design is subjective due to varying sonic preferences and perceptions among users. Considering the full range of subjective experiences is challenging. The research

did not involve user experiments, which leaves the long-term effects on user satisfaction, engagement, and perception unexplored. Longitudinal studies and monitoring of actual user behavior could provide insightful information about the long-term effects of sonification on conveying messages and on the overall experience and perception.

Given that design is a global endeavor, it would be useful to examine how different cultures interpret and perceive sounds. Understanding how cultural backgrounds and preferences influence the efficacy and acceptability of sounds can aid in the creation of culturally considerate design principles and practices.

Additionally, it would be worthwhile to explore how sounds can be combined with other sensory cues. Researching how various sensory modalities interact can help us develop a more comprehensive understanding of multisensory design principles and create more seamless and immersive experiences.

In summary, future research should address the limitations and explore the emotional impact of sounds, cross-cultural considerations, and multisensory integration. By deepening these fields, researchers can enhance the practice of sound design and create more engaging experiences. Future research should also focus on empirical testing of the developed methodology in various scenarios to validate its effectiveness and adaptability.

References

- Austin, J., (1962). *How to do things with words* (1st ed.). Harvard University Press.
- Blattner, M.M., Sumikawa, D.A. and Greenberg, R.M., (1989). Earcons and Icons: Their Structure and Common Design Principles. *Human Computer Interaction*, 4, no. 1, pp. 11–44.
- Borba (2003). *Introdução aos estudos linguísticos*. Pontes Editores.
- Brewster, S.A., Wright, P.C. and Edwards, A.D.N., (1994). A Detailed Investigation into the Effectiveness of Earcons. In: Kramer, G. (Ed.), *Auditory Display: Sonification, Audification, and Auditory Interfaces*. Reading, MA: Addison-Wesley, pp. 471–498.
- Brewster, S.A., Wright, P.C. and Edwards, A.D.N., (1995a). Experimentally Derived Guide for the Creation of Earcons. In: *Proceedings of BCS-HCI 95*. Huddersfield, UK: Springer, pp. 155–159.
- Brewster, S.A., Wright, P.C. and Edwards, A.D.N., (1995b). Parallel Earcons: Reducing the Length of Audio Messages. *International Journal of Human-Computer Studies*, 43, no. 2, pp. 153–175.
- Blauert, J. (Ed.). (2005). *Communication acoustics* (Vol. 2). Springer.
- Chandran, K.M., Prabhune, P., Sen, D. (2015). Sound Association for Product-Sound Design Using Semiotics. Chakrabarti, A. (Ed.). In *ICoRD'15 – Research into Design Across Boundaries Vol. 1. Smart Innovation, Systems and Technologies, vol 34*. (pp. 133–152). Springer. <https://DOI.org/10.1007>
- Chomsky, N. (1965). *Aspects of the theory of Syntax* (1.st edition). The Mit Press.
- Chomsky, N. (1995). *The Minimalist Program*. The Mit Press.

- Dienstfrey, E. (2016). The Myth of the Speakers: A Critical Reexamination of Dolby History. *Film History: An International Journal* 28(1), 167-193.
- Eco, H. (1999). *As Formas do Conteúdo* (3.rd.) Perspectiva.
- Eco, U. (1976). *A Theory of Semiotics* (vol. 41, no. 2). Indiana University Press.
- Eco, U. (1984). *Semiotics and the Philosophy of Language*. Indiana University Press.
- Eco, U. (2000). *Tratado Geral de Semiótica* (3rd ed.). Perspectiva.
- Farnell, A. (2010). *Designing sound*. Mit Press.
- Garner, T. A., Grimshaw, M. N., & Abdel Nabi, D. (2010). A preliminary experiment to assess the fear value of preselected sound parameters in a survival horror game. In AM '10 The 5th Audio Mostly Conference Association for Computing Machinery.
<https://doi.org/10.1145/1859799.1859809>
- Gaver, W. (1986). Auditory icons: Using sound in computer interfaces. *Human- Computer Interaction*, 2, 167-177.
- Hickey, R. (2005). *Levels of language*. Universität Duisburg-Essen.
- Jekosch, U. (2005). Assigning Meaning to Sounds - Semiotics in the Context of Product-Sound Design. In: Blauert, J. (Ed.) *Communication Acoustics*. (pp. 193-219). Springer.
https://doi.org/10.1007/3-540-27437-5_8
- Kramer G., Walker B., Bargar R., & International Community for Auditory Display. (1999). *Sonification report: status of the field and research agenda*. International Community for Auditory Display.
- Kramer, G. (1994). An introduction to auditory display. In G. Kramer (Ed.), *Auditory display: Sonification, audification, and auditory interfaces*. Reading, MA: Addison Wesley.
- Kuwano, S., Namba, S., Schick, A., Höge, H., Fastl, H., Filippou, T., Florentine, M., & Muesch, H. (2000). *The Timbre and Annoyance of Auditory Warning Signals in Different Countries*. Proc. INTERNOISE 2000, Nice, France.
- Leplâtre, G. and Brewster, S.A., 2000. *Designing Non-Speech Sounds to Support Navigation in Mobile Phone Menus*. In: Proceedings of ICAD 2000. Atlanta, GA: ICAD, pp. 190–199.
- Lopes, P., Liapis, A., & Yannakakis, G. N. (2017). Modelling affect for horror soundscapes. *IEEE Transactions on Affective Computing*, 10(2), 1-14.
- McGee, R. (2009). *Auditory displays and sonification: Introduction and overview*. University of California.
- Marcus, A. (1984). Corporate identity for iconic interface design: The graphic design perspective. *Computer Graphics and Applications*, 4(12), 24-32.

- McGookin, D. (2019). Auditory Cue Design. In M. Filimowicz (Ed.), *Foundations in Sound Design for Interactive Media* (pp. 229-258). Routledge.
- Patterson, R. (1982). *Guidelines for auditory warning systems on civil aircraft*. Civil Aviation Authority.
- Peirce, C. (1999). *Semiótica* (3rd ed.). Perspectiva.
- Salselas, I, Penha, R., Bernardes, G. (2020). *Sound design inducing attention in the context of audiovisual immersive environments*. Personal and Ubiquitous Computing.
- Saussure, F. M. (2011). *Course in general linguistics*. Columbia University Press.
- Saussure, F. (1995). *Curso de Linguística Geral* (7.^a ed.). Dom Quixote.
- Schoenberg, A. (1951). *Style and idea*. London: Williams & Norgate
- Shannon, C. (1998/1949). Communication in the presence of noise. *Proceedings of the IEEE*, 86(2), 447-457.
- Searle, J. (1969). *Speech Acts: An Essay in the Philosophy of Language*. Cambridge University Press.
- Searle, J. (1976). A Classification of Illocutionary Acts. *Language in Society*, 5(1), 1-23.
- Searle, J. (2010). *making the social world. The structure of human civilization* (1st ed.). Oxford University Press.
- Suied, C., Susini, P., Misdariis, N., Langlois, S., Smith, B. K., & McAdams, S. (2005). *Toward a sound design methodology: Application to electronic automotive sound*. Georgia Institute of Technology.
- Toprac, P. & Abdel-Meguid, A. (2010). Causing Fear, Suspense, and Anxiety Using Sound Design in Computer Games. In M. Grimshaw (xx), *Game sound technology and player interaction: Concepts and developments* (176-191). Information Science Reference.
- Wagner, A. (2006). The rules of the road, a universal visual semiotics. *International Journal for the Semiotics of Law*, 19(3), 311-3.
- Walker, B.N., & Nees, M.A. (2011). Theory of sonification. In T. Herrman, A. Hunt, & J. Neuhoff, Eds., *The Sonification Handbook* (pp. 9-39). Berlin, Germany: Logos Publishing House.