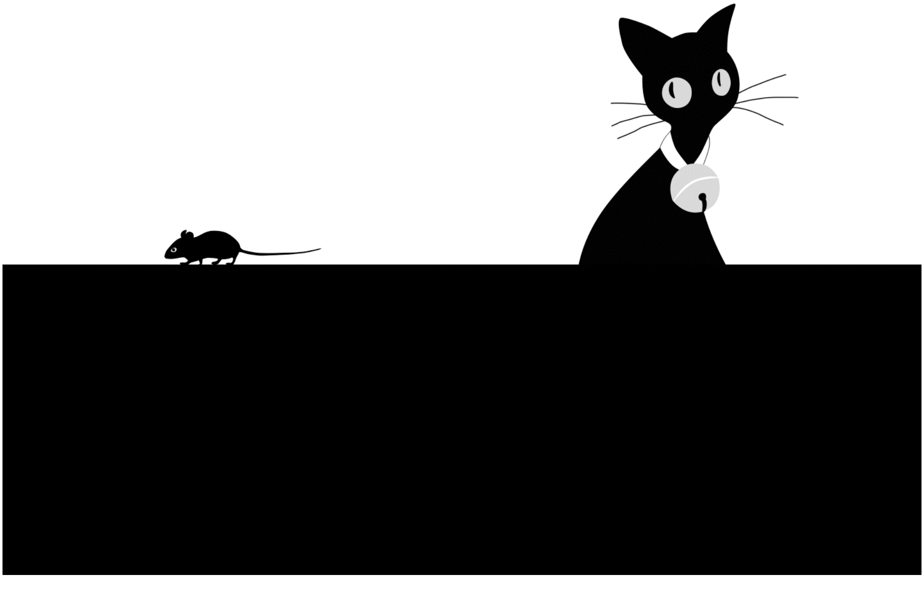
product sounds

fundamentals & application

elif özcan



ISBN 978-90-5155-048-1

product sounds

fundamentals & application

Proefschrift

ter verkrijging van de graad van doctor

aan de Technische Universiteit Delft,

op gezag van de Rector Magnificus prof.dr.ir. J.T. Fokkema

voorzitter van het College voor Promoties,

in het openbaar te verdedigen op donderdag 2 oktober 2008 om 10:00

door

Elif ÖZCAN-VIEIRA

Bachelor of Industrial Design,

Middle East Technical University, Ankara, Turkey

geboren te Ankara, Turkije

Dit proefschrift is goedgekeurd door de

Promotor
Prof. ir. J. J. Jacobs

Copromotor
Dr. R. van Egmond

Samenstelling promotiecommissie

Rector Magnificus,	Voorzitter
Prof. ir. J. J. Jacobs,	Technische Universiteit Delft, promotor
Dr. R. van Egmond,	Technische Universiteit Delft, copromotor
Prof. dr. H. de Ridder,	Technische Universiteit Delft
Prof. dr. P. Badke-Schaub,	Technische Universiteit Delft
Prof. dr. R. G. J. Meulenbroek,	Radboud Universiteit Nijmegen
Prof. dr. J. Edworthy,	University of Plymouth, UK
Dr. J. A. Ballas,	Naval Research Laboratory, USA
Prof. dr. P. P. M. Hekkert,	Technische Universiteit Delft, reservelid

“I’ve been thinking, Hobbes.”
“On a weekend?”
“Well, it wasn’t on purpose…”

~ Calvin & Hobbes

CONTENTS

⁂

Part A - Fundamentals

Part B - Application

FOREWORD

⁂

In the eye of a user, a consumer, or a potential buyer, a product can be appraised in various ways. For example, a vacuum cleaner can be powerful thus functional, a car can be serious thus trustworthy, a shaver can be immature thus unpleasant, and a computer can be elegant and most acceptable. Such appraisals determine the product-user relationship and influence decisions at purchasing or later at product usage (Desmet, 2002; Govers, 2003; Mugge, 2007; van Rompay, 2005; Sonneveld, 2007). Thus, in order to understand this relationship, both industry and academia have started to focus on developing practical knowledge as well as theories related to product experience (for an extensive review see, Schifferstein & Hekkert, 2008). In essence, these studies have demonstrated that product appraisal and a consequent meaning attribution do not necessarily originate from the product per se. A product comprises visual, auditory, tactile, and olfactory properties. It is the combination of these 'perceptual' properties and their (in)congruency with the concept of the product that give the product its deserved meaning. Hence, a narrowed down focus is needed to understand the contribution of each product property to the product experience.

This thesis particularly investigates the auditory property of products, namely product sounds. Integration of the auditory property to main design activities is a rather new topic in the industry of domestic appliances. Many, of which some successful, attempts have been made in order to do this; for example, the sounds of vacuum cleaners, coffeemakers, shavers have been considered. However, often ad hoc decisions are taken during a sound design process. In addition, (sound) designers are not supported with their sound design related activities. May this be the lack of specific tools and methodology for sound design or missing theoretical knowledge on the domain of 'product sounds' and their role in product experience.

Furthermore, designing product sounds entails an iterative exchange of expertise from various disciplines that are functionally different. In principle, the fields of acoustics, psycho-acoustics, engineering, psychology, and musicology contribute to the improvement of the sound at different stages of a sound design process. Studies regarding product sound design have often dealt with the acoustic analysis of the sound and determined their psychoacoustical correlates (see e.g., Lyon, 2001, Susini, McAdams, Winserg, Perry, Viellard, & Rodet, 2004). This method is normally used to measure the acceptability of the sound or the level of (dis)comfort the sound causes. However, product (sound) designers have come to understand that psychological effects of sounds on people cannot be restricted only to the *psychoacoustical* judgment of a sound (e.g., a sound is sharp, therefore unpleasant). Studies in auditory perception and cognition have long demonstrated that sounds that are caused by real objects and events have *meaningful associations* in memory (Ballas, 1993; Handel, 1991; McAdams, 1993; Saygin Dick, Wilson, Dronkers, & Bates, 2003). Considering that product sounds are also caused by everyday objects, the same may be true for them. Furthermore, the relation between the concept of a product and the product's sound should be congruent. This congruency can be established via conceptual associations that are common to both the product and its sound. Therefore, understanding the type of meaningful associations product sounds have will bring a new dimension to both product design and sound design.

Meaning attribution is the result of a partially perceptual and partially cognitive processing of a stimulus. Therefore, first theories and major experimental findings from the field of (auditory) cognition/perception need to be studied. Furthermore, experimental setups and methodologies used to investigate auditory cognition will be adopted for the investigation of the product sounds.

The knowledge that derives from this thesis may be of interest for designers, design researchers, and also for cognitive scientists. The expected output will be about the fundamentals of product sound perception and cognition. Designers in general and sound designers in specific could use the knowledge to understand the conceptual and physical relation between product and their sounds. Design researchers could use the knowledge to provide further theories on product (sound) experience. Because in essence, we are investigating a perceptual phenomenon of how sounds become meaningful entities rather than being treated as simply an acoustic event, the results may be of interest even to cognitive scientists.

Enjoyable reading! ☺

Elif Özcan

References

Ballas, J. A. (1993). Common factors in the identification of an assortment of brief everyday sounds. *Journal of Experimental Psychology: Human Perception and Performance, 19*(2), 250-267.

Desmet, P. M. A. (2002). *Designing Emotions*. Doctoral Dissertation, Delft University of Technology, Delft, The Netherlands.

Govers, P. (2003). *Product personality*. Doctoral Dissertation, Delft University of Technology, Delft, The Netherlands.

Handel, S. (1989). *Listening*. Massachusetts: MIT Press.

Lyon, R. H. (2000). *Designing for Product Sound Quality*. New York: Marcel Dekker, Inc.

McAdams, S. (1993). Recognition of sound sources and events. In A. S. Adams & E. Bigand (Eds.), *Thinking in sound: The cognitive psychology of human audition* (pp. 146 - 198). New York: Oxford University Press.

Mugge, R. (2007). *Product attachment*. Doctoral Dissertation, Delft University of Technology, Delft, The Netherlands.

Saygin, A. P., Dick, F., Wilson, S. M., Dronkers, N. F., & Bates, E. (2003). Neural resources for processing language and environmental sounds: evidence from aphasia. *Brain, 126*(4), 928-945.

Schifferstein H. N. J. & Hekkert, P. (2008). *Product experience*. Amsterdam: Elsevier.

Sonneveld, M. (2007). *Aesthetics of tactual experience.* Doctoral Dissertation, Delft University of Technology, Delft, The Netherlands.

Susini, P., McAdams, S., Winsberg, S., Perry, I., Viellard, S., & Rodet, X. (2004). Characterizing the sound quality of air-conditioning noise. *Applied Acoustics, 65*(8), 763-790.

Van Rompay, T. (2005). *Expressions: Embodiment in the expression of design*. Doctoral Dissertation, Delft University of Technology, Delft, The Netherlands.

INTRODUCTION

⁂

Why investigate product sounds?

The first sound a modern man hears on an average morning—even probably before the beautiful voice of a partner or happily chirping birds—is the persistent sound of an alarm clock. Our daily interaction with products and the experience of their sounds are various. One's desire for a sports car may be enhanced because of its powerful and sophisticated sound or one tends to avoid the dentist drill because of the penetrating sound. A simple microwave bell can be just as informative as a mother's call for dinner or the successive keystrokes just as confirmative as seeing the letters appearing on a screen. A ringing object in a kitchen will be treated as a kitchen timer; however, another object with almost the same sound in a bedroom will be named an old-fashioned alarm clock. In many cases of product's usage, not hearing the accustomed sound indicates malfunctioning of the product (e.g., paper jam or empty tray in a copier). The lack of sound may also result in insufficient information concerning the current status of the product usage. For example, the new owners of hybrid cars are warned to be extra cautious with unaware pedestrians while parking silent in the electric mode. Perhaps, one buys a coffee maker because it is quiet and peaceful; somebody else prefers an espresso machine perhaps because its dominating sound reminds them of happy Italian cafés.

Products are ubiquitous, so are the sounds emitted by products. Given the aforementioned examples, such sounds seem to influence our reasoning, emotional state, purchase decisions, preference, and expectations regarding the product and product's performance. Thus, auditory experience elicited by product sounds may not be just about the act of hearing or a sensory response to acoustical stimuli (e.g., this is a loud and sharp sound). People actually experience a product sound beyond its acoustical composition. People hear what the sound represents and appraise the product accordingly; or, they see what the product represents and appraise the sound accordingly. That is to say, a complimentary and meaningful relationship exists between a product and its sounds.

Industry, on the other hand, is responsible for creating these meaningful relationships when it comes to auditory ergonomics, well-being, user satisfaction, product identity, and brand differentiation. Sound is an inherent property of a product. Just like a product's visual (form, geometry, colours) and tactile (materials, texture, weight) properties, the sound also can be manipulated in order to create a desired user experience.

Given the ubiquity of product sounds and their function in our daily interaction with products, it is surprising to see that not much is known about product sounds and yet alone about how people respond to them. Much of what we know are the evident examples from the application of sound design (e.g., designed sounds of a car door or car interior). Published material on this topic often tackled sound design from an engineering and/or from a psychoacoustic point of view (Lyon, 2001; Susini, McAdams, Winsberg, Perry, Viellard, & Rodet, 2004). Available knowledge concerning experiential aspects of product sounds is limited. Some assumptions have been made to emphasize the semantic impact of sound on product experience (Jekosch, 1999; Spence & Zampini, 2006). Only few studies that dealt with product preference have provided evidence for that (Lageat, Czellar, & Laurent, 2003, Vastfjall, Gulbol, Kleiner, & Garling, 2003; Vastfjall & Kleiner, 2002). Thus, aforementioned approaches fail to provide sufficient ground to understand how people experience product sounds and what product sounds actually *mean* to them.

How can designers create a desired experience with products sounds, if they lack knowledge to predict the consequences of their decisions, if they are not supported in their conceptual thinking regarding sounds, if they fail to use a proper vocabulary that describes product sounds, and ultimately if they have no systematic methodology to design sounds? Obviously, a gap exists between the fundamentals of product sound experience and application of product sound design. This thesis bridges this gap by providing empirical findings and pointing out their relevance to the practice of product sound design.

Product sounds

Sounds emitted by products can be distinguished as *consequential* or *intentional* sounds (Van Egmond, 2008). *Consequential* sounds occur as a result of a product's functioning and its moving mechanical parts. They are mostly machinery (mechanical) sounds or sounds caused by the interaction of the user with a product.

Some examples are the hair dryer, washing machine, shaver, on-off switch, and coffee maker sounds. Due to the multiplicity of the involving parts and actions, consequential sounds produce complex sound waves. These sounds are often noisy sounds. In other words, they lack a spectral-temporal structure. Despite being noisy, consequential sounds may be informative about the state of product functioning (e.g., centrifuge cycle of a washing machine). *Intentional* sounds occur because they are chosen (often by a designer) to be a part of the product functionality or a user interface. Some examples are an alarm clock sound, a microwave oven finish signal, and feedback beeps of programming an oven. These sounds have a distinct spectral-temporal structure like musical sounds. This makes them easily distinguishable from the other environmental sounds. These sounds convey special meaning to which people attend (e.g., 'the food is ready').

In practice, sound quality assessments determine the adequacy of the sound in relation to the product (Blauert & Jekosch, 1997; see also, Fog & Pedersen, 1999). In other words, the sound should convey the same meaningful / conceptual associations as the product. Blauert and Jekosch (1997) have discussed the process by which users assess product sounds. In this process, assessment of a sound is based on auditory perception. This judgment is continuously fed by cognitive and emotional processes, and by the input from other sensory modalities. As a consequence, this framework implies that mere psychophysical measurements of a sound (e.g., sound pressure level or sharpness) or spectral analyses do not suffice to predict listeners' subjective judgments. These findings indicate the necessity of human contribution to the appraisal of sound.

Product sounds from a human perspective?

Remember the popular riddle 'if a tree falls in a forest and no one is around to hear it, does it make a sound?". This philosophical question points out to the human contribution to understanding the realm of everyday events. Events happen and they are a fact, but do they exist without perception? Products are no different within this realm. We know they exist because we see them, hear them, smell them, and sometimes taste them. We do not only perceive their existence but also infer their existence and reflect on it. Thus, investigating mere physical facts helps us understand the disposition of an object and not necessarily its mental representations.

Physical aspects of product sounds have long been investigated in the field of acoustics and engineering. In the field of acoustics, studies investigate, e.g., frequency content of a sound or its intensity in order to determine the character of a particular product sound (Letens, 2002; Susini et al., 2004). In the field of engineering, studies investigate causes of sound in order to find ways to design sounds via manipulating the product parts (e.g., Lyon, 2001) or in order to detect function failure (Benko, Petrovcic, Juricic, Tavcar, & Rejec, 2005). For determining auditory quality, engineers primarily refer to psychoacoustical judgments (sharpness, roughness, loudness, and tonalness). Although psychoacoustical responses occur on a sensory level they can be instrumentally predictable, because, people's auditory sensations are based on common physiological reactions in the ear (Bodden, 1997; see Aures, 1985 and Zwicker & Fastl, 1990 for definitions and algorithms). Determining the psychoacoustical response to a sound has been the next step engineers took to determine people’s preference for certain sounds. Nevertheless, such preference judgments do not necessarily involve meaning attribution.

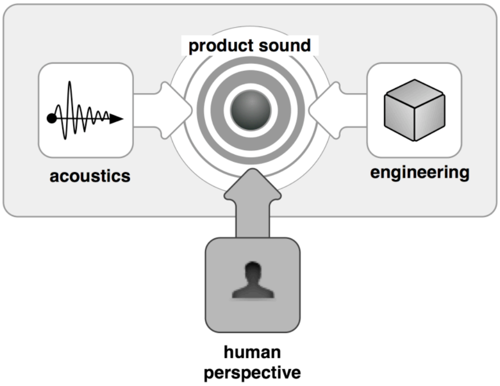


Figure 1. Old and new approaches contributing to the knowledge about product sounds.

Recent studies in product sound perception point out an experiential difference between the sensory judgments and meaning attribution. Basic sensory judgments of a product can be overridden by cognitive judgments (Lageat et al., 2003; Letens, 2002). For example, an espresso machine. The irregular, rather rough and roaring sound of a Harley Davidson can be highly appreciated by users because the sound may denote the quality of the motorbike ride (i.e., Harley as opposed to scooters) and feeling of freedom that the riding activity will bring. In another example, the loud and low-pitched sound of a vacuum cleaner suggests the powerfulness and the efficiency of the product.

Physical facts aside (namely, acoustical or engineering qualities of sound), it is the aforementioned experiential qualities that constitute the realm of product sounds in our daily lives. Thus, a new approach is necessary to discover meanings people derive from or attach to product sounds. Figure 1 shows the existing approaches that contribute to the knowledge we have about product sounds. It also indicates the addition of a new approach from human perspective. Ultimately, all the motives to design product sounds concern people (buyers, consumers, users, and sometimes designers) and people's behavioural tendencies towards products. Thus, understanding the human perspective on sounds will make the existing knowledge more relevant to people's experiences.

Product sounds within the network of conceptual associations

Murphy (2002) describes concepts as "the glue that holds our mental world together". Imagine a shaver. Our knowledge about this particular product will contain information about the function of the product ('personal hygiene' and 'shaving') and where it belongs (bathroom), about how it looks (round edges, black, metallic colour), how it feels in the hands (plastic, soft texture) and eventually how it sounds (loud, high-pitched). Furthermore, these properties of a product all together convey higher-level associations, such as a *futuristic look,* and *expensiveness*. All this knowledge is glued together by the *concept* of a shaver. Seeing, hearing, feeling a product, interacting with it, or being in a certain location will activate a bundle of relevant information that is glued together by concepts (Bartlett, 1977; Paivio, 1991; Thompson & Paivio, 1994). Accessing concepts, therefore, is a fundamental cognitive action in meaning attribution.

What constitutes concepts? In memory, information about an object is contained in modality-specific sensory memories and in a semantic memory (see, e.g., Paivio, 1991; Barsalou, 1999). **Sensory memories** contain perceptual information regarding the physical properties of an object. For example, a visual property (shape, colour) of an object is stored in the *visual memory*, and an auditory property (spectral-temporal content) in the *auditory memory*. **Semantic memory** contains verbal information corresponding to sensory memories and also to concepts. In the semantic memory, concepts can be described concisely by labels, which are also referred to as *lexical representations.* Thus, accessing a lexical representation in memory would be satisfactory to activate a network of conceptual associations.

Considering the constitution of concepts, it seems impossible to isolate meanings attached to sounds from the influence of other product properties. Therefore, in this thesis product sounds will be investigated through the concept of a product. Figure 2 presents the focus of this thesis. In the figure, product properties are divided into three: visual, auditory, and semantic. In memory, these properties are continuously in interaction with each other via the product concept. Accordingly, meanings derived from the auditory property of the product may be subject to changes depending on the influence of visual and semantic properties. Furthermore, the effect of context in meaning attribution should also be considered, because many products are location-specific and thus may be more meaningful within context.

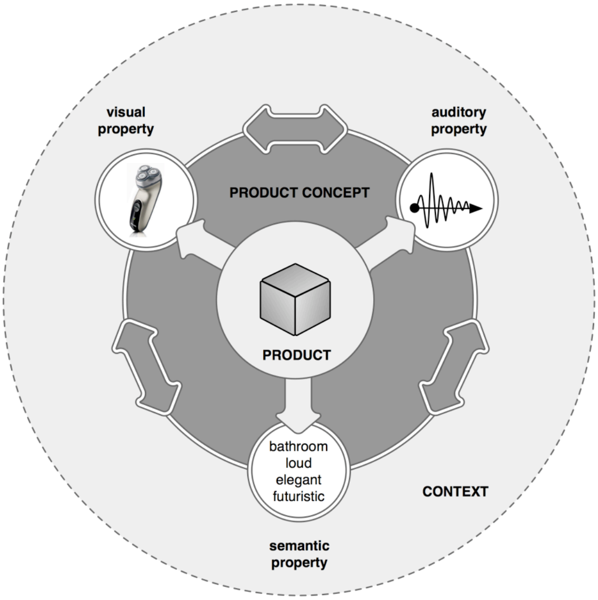


Figure 2. Sound as a property of a product.

Meaning attribution

In psychological terms, meaning attribution occurs during the process of object identification. The process of identification may contain a variety of sub-processes such as perceptual analysis, recognition, and identification (see Figure 3). This complex process requires the co-operation of various cognitive functions at different stages of identifying an object (Biederman, 1987; Handel, 1991; Stevenson & Boakes, 2003). ***Memory*** plays an important role at acquisition of information, storing, recognizing, and recalling. Different memory systems interact with each other. Thus, recognizing the perceptual attributes of an object may lead to access a semantic, or a lexical store (recalling the name). During recognition, perceived structural features of an object are mapped onto previously coded structural features of an object. A ***categorical judgment*** can be made upon recognition and be dependent on ***similarity judgments***. Assigning a category may activate all conceptual associations concerning an object. Once an object is recognized and categorized, access to a semantic and to a lexical store may become easier. ***Semantic associations*** of the percept can occur at this stage that results in conceptual identification. However, ideally identification occurs when a ***lexical representation*** is provided in the form of an object name. It is possible that although recognition has occurred, access to a lexical store fails, or multiple lexical associations occur. This may cause ***ambiguity*** in identification. Ambiguity may hinder the completion of the identification process. In such cases, the ***context*** in which the object is presented may facilitate identification. For sound identification similar processes and cognitive functions apply (Bregman, 1990; Handel, 1989; McAdams, 1993).

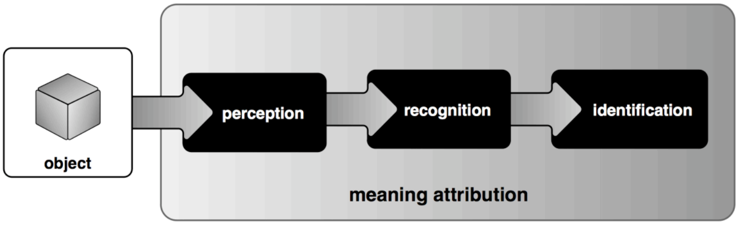


Figure 3. Process of meaning attribution.

Thesis objectives

The primary questions that concern this thesis are: What actually underlies people's experience with product sounds? What is the relationship between a product and its sound? Are there any external factors that might influence this relationship? Answers ultimately concern the design practice. With the answers provided by this thesis, design practitioners and scholars will gain knowledge on the psychological effects of sounds on people, and consequently, on product appraisal.

This thesis aims to determine the conceptual and semantic networks for product sounds by investigating the perceptual and cognitive processes that result in sound identification. In this Introduction, sound identification has been discussed in terms of its constituting basic cognitive functions and in terms of resulting semantic associations. The remaining chapters will focus in detail on each of these cognitive functions. Gradually, a network will be built that organizes the conceptual associations of product sounds and determine factors that may influence the organization of the semantic knowledge within this network.

Ultimately, the thesis aims to draw attention to product sound design as an upcoming discipline. Accordingly, methods will be proposed to support designers/engineers in their sound design activities. Responsibilities of sound designers will be determined with respect to the multi-disciplinary nature of a sound design task. Finally, the plausibility of the product sound design as an independent discipline will be discussed.

This thesis

This thesis contains two main parts (see Figure 4). Part A concerns the fundamentals of product sound cognition. Part B concerns the application of product sound design. **Part A** consists of experimental studies and theoretical findings related to object perception/identification and their relevance to product sound identification. Because not much is known about product sounds, Chapters 1 and 2 contain explorative studies. These studies analyze listeners' direct responses to product sounds using free categorization and free labeling paradigms. After collecting the preliminary information about how sounds are mentally represented, in Chapters 3, 4, and 5, the validity of these findings are tested and more insight into external factors that intervene with the sound identification process are gained. **Part B** consists of a review and discussion of the existing knowledge on the practice of sound design. Figure 4 summarizes the conducted studies and indicates the focus of each chapter.

Part A - Fundamentals

Regarding Part A, Figure 4 presents a conceptual framework that demonstrates the stages of a meaning attribution process. In this framework, listeners' responses to sound are examined at different stages of an identification process (perception, recognition, and identification). In addition, a sound is considered to have a conceptual relationship with the other product properties (visual and semantic). Therefore, the influence of other product properties is investigated on different stages of the sound identification process.

In **Chapter 1**, the domain of product sounds is determined in terms of perceptual categories through similarity judgments. Basic concepts that represent the categories, the categories' relevant semantic associations and underlying psychoacoustical correlates are defined. In **Chapter 2**, the emerging concepts and semantic associations are specifically investigated for each product sound category that has been determined in Chapter 1. Characterizing sound descriptions are determined. In **Chapter 3**, memory for product sounds is investigated through semantic priming / encoding methods (self labeling, verbal and visual labels). In **Chapter 4**, the extent to which product sounds are identified are investigated. In addition, whether ambiguity is a factor that influences the ability to correctly identify a product sounds is discussed. In **Chapter 5**, the (positive) effect of visual context on the identification of ambiguous product sounds will be investigated. The visual information varies in the degree of semantic information. **Chapters 3 & 5** also investigate the interconnections between auditory and visual memory for product sounds and the extent to which visual information has an additive effect on sound identification.

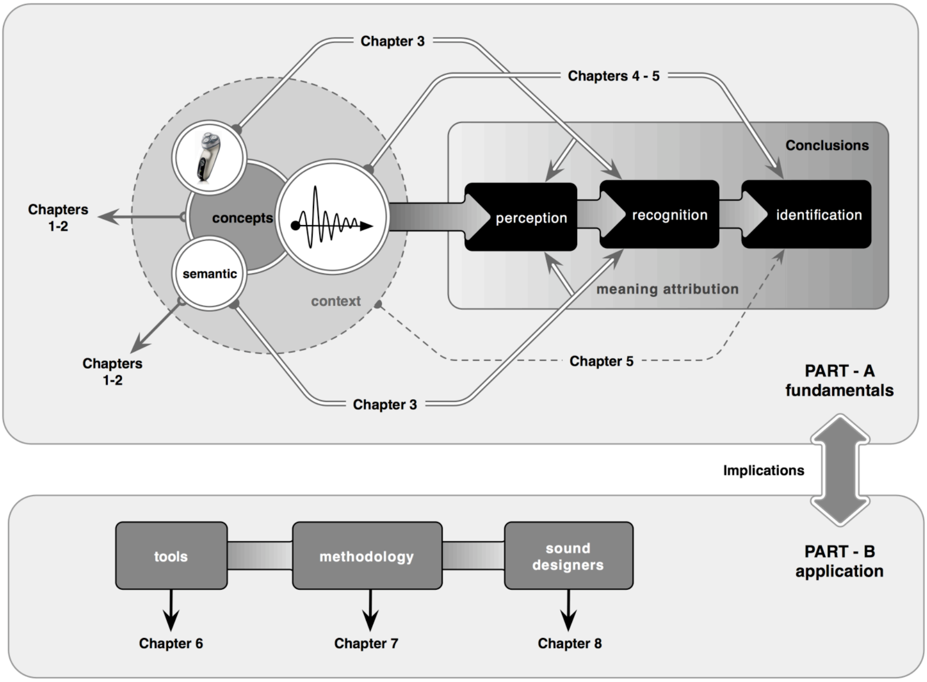


Figure 4. A summary of the studies with respect to the chapters they occur.

Finally, in the **Conclusions** section of Part A, the empirical findings are summarized in a framework that describes the proposed process of meaning attribution for product sounds.

Part B – Application

In Part B, product sound design as an upcoming discipline will be discussed. Therefore, **Chapter 6** proposes a new visual tool that can facilitate the communication of sound characteristics during a design activity. **Chapter 7** reviews existing methods of product development and proposes a new methodology for designing product sounds. **Chapter 8** analyzes the disciplines contributing to product sound design and points out the responsibilities of a sound designer.

The **Implications** section will make the experimental findings relevant to the application of product sound design. Suggestions for future studies are also discussed.

Note that this thesis is a conglomeration of published / submitted papers. A theoretical background has been provided for each chapter. Readers can read each chapter independently of others.

References

Aures, W. (1985). Berechnungsverfahren für den sensorischen Wohlklang beliebiger Schallsignale. *Acustica*, *59*,130-141.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22*(4), 577-660.

Bartlett, J. C. (1977). Remembering environmental sounds: Role of verbalization at input. *Memory & Cognition, 5*(4), 404-414.

Benko, U., Petrovcic, J., Juricic, D., Tavcar, J., & Rejec, J. (2005). An approach to fault diagnosis of vacuum cleaner motors based on sound analysis. *Mechanical Systems and Signal Processing, 19*(2), 427-445.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review, 94*(2), 115-147.

Blauert, J., & Jekosch, U. (1997). Sound-quality evaluation - A multi-layered problem. *Acustica, 83*(5), 747-753.

Bodden, M. (1997). Instrumentation for sound quality evaluation. *Acustica, 83,* 775-783.

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound* Massachusetts: MIT Press.

Fog, L., & Pedersen, T. (1999). Tools for product optimization. *Human Centered Processes 10th Mini EURO Conference*, Brest, France.

Handel, S. (1989). *Listening*. Massachusetts: MIT Press.

Jekosch, U. (1999). Meaning in the context of sound quality assessment. *Acustica, 85*(5), 681-684.

Lageat, T., Czellar, S., & Laurent, G. (2003). Engineering hedonic attributes to generate perceptions of luxury: Consumer perception of an everyday sound. *Marketing Letters, 14*(2), 97-109.

Letens, U. (2000). *Exemplarische Darstellung einer subjektiven Beurteilung von PKW-Fahrgeräuschen.* Paper presented at the Fortschritte der Akustik DAGA 2000, Oldenburg.

Lyon, R. H. (2000). *Designing for Product Sound Quality*. New York: Marcel Dekker, Inc.

McAdams, S. (1993). Recognition of sound sources and events. In A. S. Adams & E. Bigand (Eds.), *Thinking in sound: The cognitive psychology of human audition* (pp. 146 - 198). New York: Oxford University Press.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology - Revue Canadienne De Psychologie, 45*(3), 255-287.

Spence, C., & Zampini, M. (2006). Auditory contributions to multi-sensory product perception. *Acta Acustica, 92*(6), 1009-1025.

Stevenson, R. J., & Boakes, R. A. (2003). A mnemonic theory of odor perception. *Psychological Review, 110*(2), 340-364.

Susini, P., McAdams, S., Winsberg, S., Perry, I., Viellard, S., & Rodet, X. (2004). Characterizing the sound quality of air-conditioning noise. *Applied Acoustics, 65*(8), 763-790.

Thompson, V. A., & Paivio, A. (1994). Memory for pictures and sounds: independence of auditory and visual codes. *Canadian Journal of Experimental Psychology, 48*(3), 380-398.

Van Egmond, R. (2008). The experience of product sounds. In H. N. J. Schifferstein & P. Hekkert (Eds.), *Product experience*. Amsterdam: Elsevier.

Vastfjall, D., Gulbol, M. A., Kleiner, M., & Garling, T. (2002). Affective evaluations of and reactions to exterior and interior vehicle auditory quality. *Journal of Sound and Vibration, 255*(3), 501-518.

Vastfjall, D., & Kleiner, M. (2002). *Emotion in product sound design.* Paper presented at the Journees Design Sonore.

Zwicker, E., & Fastl, H. (1990). *Psychoacoustics: Facts and models*. Berlin: Springer.

PART A

⁂

FUNDAMENTALS

This chapter is based on the paper:

Özcan, E. van Egmond, R., & Jacobs, J. (2008) Categorization and identification of product sounds. Manuscript submitted to *Journal of Experimental Psychology: Applied.*

Abstract

In five experiments, the categorization of and the conceptual and semantic associations of product sounds were investigated. In Experiment 1, main product sound categories were obtained and basic concepts were determined that categorize the mental representation of the occurring sound categories. In Experiment 2, the bases for the categorization of very similar sounds were determined. Experiments 3 and 4 separately investigated the semantic associations of individual product sounds and product sound categories. In Experiment 5, strategies for similarity judgments of product sounds were determined. As a result, this study provided six perceptually distinct product sound categories and conceptual associations that distinguish between these categories. These conceptual associations seem to result mainly from perceptual and cognitive evaluations of a product sound. Consequently, the semantic associations of a product sound are dependent on at which stage of the identification process a sound is identified.

CHAPTER 1

⁂

BASES FOR CATEGORIZATION AND IDENTIFICATION OF PRODUCT SOUNDS

Domestic appliances (i.e., products) are being used in order to facilitate our modern life style. Waking up by the sound of an alarm clock, using an electrical toothbrush to have clean teeth, and preparing a cup of espresso are a few examples of the daily activities in which products are involved. Most products consist of functional parts that, when energy—an electrical or a manual source—is applied, start moving and consequently produce sounds. These consequentially occurring sounds inform us about the functioning of a product, but also produce affective reactions, and influence reasoning, purchasing decisions, and preferences regarding the product (Lageat, Czellar, & Laurant, 2003; Spence & Zampini, 2006; Västfjäll, Kleiner, & Gärling, 2003a, 2003b; Zampini, Guest, & Spence, 2003). These findings also imply that the auditory experience of products is often determined by the meaningful and conceptual relationship between a product and the sound it produces. Although much is known about the processes of meaning attribution to the visual features of a product (e.g., Biederman, 1987; Palmer, 1975; van Rompay, 2008; van Rompay, Hekkert, Saakes, & Russo, 2005), knowledge on how people perceive product sounds is limited (see Lyon, 2000; Özcan & van Egmond, 2006). Studies that have investigated the perception of environmental sounds also contained product sounds (e.g., Aldrich, 2005; Ballas, 1993; Kidd & Watson, 2003). However, these sounds have not been explicitly studied as a separate perceptual domain. Therefore, in this study, the categorization of product sounds their conceptual and semantic associations will be investigated. First, the domain of environmental sounds will be discussed.

Environmental sound domain

Environmental sounds are considered to be meaningful because they are generated by real objects and events (e.g., Saygin, Dick, & Bates, 2001). Furthermore, there are empirical findings that such sounds have a conceptual relationship to their sources and to the context in which they occur (Aldrich, 2005; Ballas 1993; Ballas & Mullins, 1991; Bartlett, 1977). The verbal attributes that describe environmental sounds can range from the description of the perceptual features of the sound to the description of its source. (Björk, 1985; Fabiani, Kazmerski, Cycowicz, & Friedman, 1996; Kidd & Watson, 2003; Marcell, Borella, Greene, Kerr, & Rogers, 2000). According to Vanderveer (1979), description of the auditory features occurs because the sound source has not been identified. Thus, semantic and conceptual associations of environmental sounds may occur at different levels as a result of the level of identification.

Memory performance for environmental sounds also benefits from (visual or verbal) labeling. Remembering sounds is easier if they are encoded with a label (Bartlett, 1977). Interconnections between an auditory store and a visual store create an additive effect on the memory for environmental sounds (Özcan & van Egmond, 2007; Thompson & Paivio, 1994). Moreover, available structure at encoding facilitates the retrieval of auditory codes (Deutsch, 1980; Deutsch & Feroe, 1981; Özcan & van Egmond, 2007).

The factors that affect the identifiability of environmental sounds may vary. Bottom-up processes require structure in the spectral-temporal content of the sound for storing auditory information and retrieving its meaning (Bregman, 1990; Gygi, Kidd, & Watson, 2004; McAdams, 1993; Warren, 1993). Top-down processes may benefit, e.g., from the effect of context in which the sound most likely occurs (Ballas & Mullins, 1991). Yost (1991) as well as Ballas (1993) have suggested that environmental sound identification occurs via a process that incorporates perceptual and cognitive analyses. Yost (1991) has emphasized the importance of auditory imagery in sound identification (i.e., ability to associate the sound with a source or with a perceptual category). However, he has also suggested that a lexical identification of the sound source is not obligatory for sound identification and thus, sound source determination and identification are separate processes. In addition to this, Ballas (1993) emphasized the importance of familiarity, ease of naming (i.e., lexical associations), ambiguity (i.e., assessment of the alternative causes), and ecological relevance (i.e., context to which sound sources belong) in sound identification.

Product sounds

Product sounds need to be investigated as a separate domain because understanding them may help (design) engineers enhance the quality of the auditory experience and consequently the product experience (van Egmond, 2008). Most studies that have investigated product sounds have used a technical approach to understand them (e.g., spectral content, sound quality). Often the acoustical character of such sounds has been investigated (Benko, Petrovcic, Juricic, Tavcar, Rejec, & Stefanovska, 2004; Susini, McAdams, Winsberg, Perry, Viellard, & Rodet, 2004).

The contribution of audition to product experience can be both on a semantic level and on an emotional level (Spence & Zampini, 2006; Västfjäll, Gulbol, Kleiner, & Gärling, 2002). The emotional impact of a sound may result from a sensory evaluation (e.g., sharpness) or a cognitive judgment (e.g., unpleasant). Semantic judgments on a sound may determine the functional use of the sound (and the product). Semantic or emotional judgments have implications for the acceptability of the product (Susini et al., 2004, Lageat at al., 2003). Therefore, product sounds can be judged on several verbal attributes that relate to the product emitting the sound (Blauert & Jekosch, 1997; Bisping, 1997). In addition to sounds produced by the operating or machinery of a product, there are sounds that are implemented in the products (e.g., alarm and feedback sounds). These sounds are more abstract but yet meaningful and often their meaning needs to be learned (Edworthy, Hellier, & Hards, 1995).

The production of all sounds—except synthesized warning signals—is based on a non-arbitrary set of relationships between the sound source in action and the acoustic outcome of the event. Consequently, if the signal-referent associations are strong, a sound can communicate about an event in terms of its function (Keller & Stevens, 2004; Petocz, Keller, & Stevens, in press; see also Jekosch, 1999). Such strong associations may be established quickly when the event is explicit to a listener. However, not all product sounds may have such communicative value, because many product sounds are actually a result of mechanisms that are invisible to a listener. Meaning attribution to such sounds may operate via learned associations to the source information. Therefore, mental representations of a product sound may also contain source related information.

Mental representations

Theories regarding object representation in memory basically discern three types of mental representations: modality-specific perceptual symbols, concepts, and semantic knowledge (see, e.g., Barsalou, 1999; Paivio, 1991). Each perceptual system holds modality specific information that is linked to meaningful associations such as concepts and semantic knowledge. These representations co-operate in the process of recognizing an object and finally identifying it. Because our focus is on meaning attribution, we will discuss the differences and commonalities between *conceptual representations* and *semantic representations*.

Conceptual representations

According to Te Linde (1983) and Paivio (1991), a *concept* is the common space for the perceptual and the corresponding semantic knowledge. That is, perceptual knowledge concerning an object (e.g., shape, colour, frequency content, temperature) is processed by a non-verbal system and stored in modality specific systems. Semantic correspondence of an object is processed and stored by a verbal system. Pecher, Zeelenberg, and Barsalou (2003) regard the information from different modalities as the properties that constitute the concept. Concepts undergo a variety of cognitive functions such as recognition, categorization, and identification (Solomon, Medin, & Lynch, 1999). Categories and concepts are often treated the same. However, Murphy (2002) states that a concept is the mental representation of classes of objects, whereas *categories* are the classes themselves.

Rosch (1978) and Mervis and Rosch (1981) have suggested that a *category* exists whenever two or more distinguishable objects are treated equivalently. They distinguish between a vertical organization that is dependent on the level of specificity and a horizontal organization depending on the similarity between members. On the horizontal level a category member that is the most similar to the other members is called the ‘typical member’ that represents a category. On the vertical level, an object can be hierarchically represented, for example, it can belong to the *chair* category, *dining chair* category, or *furniture* category. Many studies (Borghi, Caramelli, & Setti, 2005; Murphy & Smith, 1982; Rosch, 1978; Tversky & Hemenway, 1984) have investigated the specificity of the perceptual and conceptual information conveyed by such categories. On a basic level, objects are similar in terms of parts and actions (e.g., birds have feathers and wings and they can fly). On a sub-ordinate level, perceptual and featural information are necessary to distinguish between objects that eventually belong to the same basic level category (e.g., comparison of robin vs. sparrow in the bird category). On a super-ordinate level, more conceptual information (e.g., functions, context) is required to distinguish between objects or concepts (furniture vs. animals).

Similarity judgments for categorization can be a result of a perceptual process or a cognitive activity. For visual similarity judgments, Eme and Marquer (1998) have demonstrated that, different individuals may spontaneously adopt different strategies and an individual may apply different strategies on different trials (e.g., holistic, analytic, one-feature). Medin and Barsalou (1987) have suggested that categories occur as a result of sensory perception or generic knowledge. *Sensory perception* categories stem from the similarities in the perceptual features (e.g., colour, sound). *Generic knowledge* categories stem from the similarities in the conceptual knowledge (e.g., birds, cars) that is linked to meaningful semantic associations (e.g., birds have wings and they can fly).

Theoretically, the type of event causing the sound inherently determines the acoustical character of a sound. For example, Gaver (1993b) has discerned impact (solids), scraping (solids), dripping (liquids), temporally complex events (interaction of solids and liquids), and machine sounds (complex events involving various sources) within the environmental sound domain. From an empirical account, the methods of sound categorization have been based on people's observations and subjective sound descriptions (Porteous & Mastin, 1985), on cognitive judgments (Ballas, 1993), and on the first conceptual representation that is activated upon the auditory perception (Marcell et al., 2000). Consequently, sound categories reflect varying degrees of common conceptual knowledge (e.g., nature, bathroom, water, door, indicator / signaling sounds) and somewhat common perceptual features (modulated noise, sounds with two to three transient components).

The paradigm used may also moderate the type of similarity judgments (Aldrich, Hellier, & Edworthy, in press; Gygi, Kidd, & Watson, 2007). For example, paired comparison paradigms may result in categories of sounds that have common spectral-temporal composition; free-categorization paradigms may require the cognitive analysis of the sounds. However, if sound categories are the result of signal-referent relationship, then both acoustical and conceptual associations could result in similar categories. Moreover, when Gaver's (1993b) and Ballas's (1993) categories are compared, one can see overlaps although one type of categories resulted from acoustical estimates of interacting materials and the other from cognitive judgments (e.g., common water sound categories, vibrating solids category of Gaver as opposed to the door and engine sounds of Ballas).

The similarity of imagined sounds and the similarity of imagined sound sources are negatively correlated (Gygi et al., 2007). This may indicate that categorical and semantic arrangements of sounds do not require acoustical similarity. Nevertheless if the signal-referent relationship is strong, then conceptual and perceptual categories should resemble each other. Otherwise, the observed salient similarity between the available sounds should determine the categorical consistency, because a sound may belong to multiple categories (Marcell et al., 2000).

Semantic representations

Concepts are the bridge between the sensory specific information and the semantic knowledge. Semantic associations are the verbal correspondents of what a concept represents. Because a concept can hold information from different modalities, relevant semantic associations can be easily activated when a concept is activated. Consequently, conceptually identifying an object may also allow one to describe the perceptual features of the object, and what the object represents. For example, the semantic associations of a hairdryer concept could be loud sound, air blowing, warm, styling, drying, bathroom. A lexical representation is a specific type of semantic association that can solely represent a concept with as few words as possible (e.g., hairdryer, Braun hairdryer, my hairdryer). A lexical representation can be considered a compact yet meaningful summary of a concept. Lexically identifying an object will activate the relevant semantic associations.

Lexical representation of a sound is often determined by the sound source and the action causing the sound (e.g., car passing) (see Fabiani et al., 1996; Marcell et al. 2000; Vanderveer, 1979). Semantic associations are activated earlier than the lexical associations (Cummings, Ceponiene, Koyama, Saygin, Townsend, & Dick, 2006; McCauley, Parmelee, Sperber, & Carr, 1980; see also Cleary, 2002). In other words, people are able to verbalize their auditory percept before a complete identification occurs. Moreover, if people fail to identify or to access any lexical representations, they are still able to verbalize the psycho-acoustical and structural properties of the sounds (Vanderveer, 1979). Fabiani et al. (1996) have categorized semantic and lexical representations of environmental sounds as not-known (e.g., noise), sound imitation (e.g., too-too-too), sound description (e.g., high-pitched), name or compound name (e.g., bird, water drain bubbles). In this study, the level of the conceptual representation was also determined: car for *modal;* automobile for *synonym;* truck for *coordinate;* vehicle for *super-ordinate;* Ferrari for *sub-ordinate.*

Synthesized sounds also elicit semantic associations (Edworthy et al., 1995; Solomon; 1958; von Bismarck, 1974). These associations refer to abstract concepts (e.g., controlled, dangerous, steady, urgent, etc.), sensory experience (e.g., unpleasant, obtrusive), and/or psychoacoustical character of the sound (e.g., sharp, high, loud). Changes in the acoustic dimensions (e. g., pitch, speed, inharmonicity, and rhythm) influence the perceived meaning of an abstract sound (Edworthy et al., 1995).

In the current study, the relation between the acoustical information and relevant perceptual features (e.g., spectral temporal content) will be determined. In addition, the semantic associations and basic concepts that characterize product sounds will be determined. To this end, five experiments have been conducted to provide further insight into the ways listeners perceive product sounds. Experiment 1 investigates the product sound categories and the underlying dimensions. Experiment 2 investigates the underlying factors for listeners' categorization of (acoustically) similar product sounds. In Experiment 3 and 4, the factors that determine the semantic associations of product sounds are investigated. Experiment 5 determines on which level (e.g., perceptual or cognitive) listeners' similarity judgments occur. The sounds used in the experiments stem from commonly used domestic appliances.

Experiment 1

Categorization of product sounds

Studies that investigated categories in the environmental sound domain have often focused on the semantic relation of the category members rather than their similarity based on the perceptual features (e.g., Ballas, 1993; Marcell et al., 2000). A direct comparison between sounds is necessary to determine category borders and consistent memberships. The two main objectives of this experiment are to determine (1) the domain of domestic product sounds together with its constituting categories, (2) the acoustical / psychoacoustical dimensions that underlie this domain. The experiment consists of four tasks. First, product sounds were grouped on the basis of their perceptual similarity using a free categorization task. Secondly, each group was labeled. Thirdly, the fit of the sounds in a category was rated. Fourthly, the most representative sound within a category was determined.

Method

Participants

Twenty-eight students and employees of Delft University of Technology volunteered. The average age of the participants was 27.5 years. All participants reported normal hearing.

Stimuli

Thirty-two domestic product sounds were selected from various sound effect CDs. The sounds were edited on a Macintosh PowerPC G4 computer using the sound-editing program Sound Studio. Sounds longer than 5 seconds were trimmed to a maximum duration of 5 seconds. Sounds that were shorter than 5 seconds were not changed in duration. All sounds were saved in a stereo format with a sampling rate of 44.1 kHz and 16 bits. The loudness levels were adjusted to a comfortable listening level for each sound. The participants were not allowed to change the sound levels during the experiment. Table 1 presents the psycho-acoustical parameters for each recorded sound. The sound pressure level (SPL) of each sound was measured by a Bruel & Kjaer 2260 Investigator in a quiet room. The SPL level was used to calibrate

			Psychoacoustical metrics					
Category	**ID**	**Sound Description**	**SPL**	**SC**	**S(Z&F)**	**S(A)**	**Int. (E-05)**	**Harm.**
Air	1	centrifuge cycle (WM)	78	1697	1.8	5.1	6.30	-1.35
	2	hairdryer	75	3790	2.0	5.7	3.20	-1.57
	3	vacuum cleaner	76	889	1.3	3.4	4.00	-0.04
	4	vacuum cleaner (hand)	74	1671	1.6	4.1	2.50	-1.79
	5	washing machine	69	1673	1.7	4.2	0.79	-2.02
		Mean	*74*	*1944*	*1.7*	*4.5*	*3.35*	*-1.35*
Alarm	6	alarm clock (digital)	79	5471	2.4	6.5	7.90	17.41
	7	finish beep (MWO)	65	2321	1.9	3.7	0.32	27.97
	8	finish bell (MWO)	65	8670	2.2	4.5	0.32	14.07
	9	setting (MWO)	63	2144	1.5	2.7	0.20	20.54
		Mean	*68*	*4652*	*2.0*	*4.4*	*2.19*	*20.00*
Cylic	10	dishwasher	70	272	1.3	2.7	1.00	-1.49
	11	kitchen extractor fan	75	681	1.4	3.7	3.20	-3.66
	12	microwave oven	73	267	1.2	2.5	2.00	0.5
	13	tumble dryer	76	234	1.3	3.0	4.00	1.95
		Mean	*74*	*364*	*1.3*	*3.0*	*2.53*	*-0.67*
Impact	14	program selection (TD)	65	1302	1.5	3.2	0.32	-2.53
	15	door closing (MWO)	78	-	-	-	6.30	-1.36
	16	door closing (WM)	70	-	-	-	1.00	-5.34
	17	door opening (MWO)	77	-	-	-	5.00	-1.91
	18	door opening (WM)	76	-	-	-	4.00	-4.88
	19	on-off button (KEF)	75	-	-	-	3.20	-4.97
	20	on-off button (MWO)	77	-	-	-	5.00	3.47
	21	on-off button (ventilator)	74	-	-	-	2.50	-2.66
	22	on-off button (WM)	69	-	-	-	0.79	-4.75
	23	popping up toast (toaster)	77	-	-	-	5.00	-3.03
	24	nail click (SM)	74	-	-	-	2.50	1.33
		Mean	*74*	*1302*	*1.5*	*3.2*	*3.24*	*-2.42*
Liquid	25	coffee boiling (CM)	73	856	1.4	3.2	2.00	-3.11
	26	coffee brewing (CM)	68	1407	1.5	3.5	0.63	4.71
	27	water boiling (kettle)	74	439	1.0	2.1	2.50	-2.07
	28	water pouring (CM)	66	2748	1.8	4.2	0.40	-4.32
		Mean	*70*	*1363*	*1.4*	*3.3*	*1.38*	*-1.20*
Mechanical	29	alarm clock (mechanical)	79	7671	2.5	7.4	7.90	-4.87
	30	shaver	74	2584	2.0	5.2	2.50	12.7
	31	toothbrush	71	3341	2.2	5.6	1.30	0.68
	32	winding (phonograph)	72	1949	1.6	4.0	1.60	-1.02

Mean	*74*	*3886*	*2.1*	*5.6*	*3.32*	*1.87*

Note. ID = sound number used in Figure 1, SPL = sound pressure level measured in decibels, SC = spectral centroid, S(Z&F) = sharpness algorithm defined by Zwicker & Fastl (1993), S(A) = sharpness algorithm defined by Aures (1985), Int. = sound intensity (in W/m^2), Harm. = harmonicity. WM = washing machine, MWO = microwave oven, TD = tumble dryer, KEF = kitchen extractor fan, SM = sewing machine, CM = coffee machine.

Table 1. Psychoacoustical metrics calculated for each product sound and sound category.

the psycho-acoustical analysis software. Harmonicity was calculated using Praat[1]. Two sharpness parameters (Zwicker & Fastl, 1990; Aures, 1985), the spectral centroid, intensity, and 39 critical band levels in Erbs were calculated using Psysound[2]. To reduce the number of parameters a principal component analysis (PCA) with Varimax rotation was conducted that resulted in two factors explaining 74% of the variance. High frequency critical bands (CB-1747 through CB-15085), the two sharpness parameters, the spectral centroid, the SPL, and intensity level loaded high and positively on *Factor 1* (explaining 44.70% of the variance). Therefore, this factor was interpreted as a combination of sharpness and loudness. Low frequency critical bands (CB-55 through CB-1545) loaded high and positively on *Factor 2*, whereas Harmonicity loaded high and negatively on *Factor 2* (explaining 33.73 of the variance). Therefore, this factor was interpreted as a combination of low frequencies and noisiness. In addition, the regression weights for *Factor 1 and Factor 2* were extracted for each sound.

Apparatus

The stimuli were presented using a specially designed software program developed with the Trolltech Qt (Mac OS X - free edition) tool kit. The program ran on a Macintosh Power PC G4 computer with a 1s7” Iiyama Pro454 monitor. Apple Pro Speakers with a frequency range of 70Hz - 20kHz and a maximum power of 10 Watts per channel were used to present the stimuli. The study took place in a quiet room.

Procedure

Before the study started, each participant received a brief explanation about the purpose of the study on an A4 sized paper. A free categorization paradigm was used. That is, a participant's task was to freely group the sounds they considered similar.

Prior to the actual experimental session, a participant took a training session with animal sounds and human voices. The tasks in the training session were identical to the experiment, only the stimuli differed. In the experimental session, the sounds were presented as buttons on the computer screen. The sound buttons were divided into two functional parts. The part with the number was used to drag the sound button. The part with the small speaker icon was used to listen to the sound. A button remained dimmed until a participant listened to the sound. For every participant the numbers were randomly assigned to a sound. A participant had to listen to all the sounds and freely group them on the screen. After a participant heard all the sounds, s/he could advance to the next stage. In this stage, a participant had to create 'boxes' in which the previously defined groups could be dragged. Each created box (i.e., sound group) had to be labeled by the participant. This label had to reflect how they would describe the group. Note that no instruction was given what type of label a listener had to give (e.g., source, interaction event, or emotional experience, etc.). A participant set the degree of fit on a 7-point scale ranging from one to seven (how well do the sounds fit together?) for each group. In addition, a participant chose the most representative product sound in each group. A participant received a warning on the screen for each step s/he failed to progress. After the categorization task, a participant was debriefed to understand the associations made between sounds and groups.

Results

The minimum number of categories created was two and the maximum number of categories was nine in the grouping task. Sixteen participants created five categories. The mean for the category fit ratings was 4.98 and differed significantly from the middle-point of the scale (*t(27)=6.96, p<.001*). This indicated that on average the participants were satisfied with their groups. A multidimensional scaling technique was used to determine the categories. Individual category data were transformed into a matrix consisting of a dummy coding (0 and 1 values). The Proxscal procedure (SPSS) was used to analyze these multiple matrix sources. Proxscal transformed the counts into a Chi-square measure. The distance matrix was then scaled using the Identity model into a forced 2-dimensional solution that yielded coordinates for each sound. The 2-dimensional solution had a Stress-I value of .08. The 3-dimensional solution was not used because it resulted in only a minor decrease in stress (from .08 to .07). In addition, the effect of sample size on the solution was tested. Three sets (*n*=14) were randomly drawn from the total dataset (*N*=28) and were analyzed using the aforementioned MDS procedure. The correlations between the coordinates of these three sets and the coordinates of the entire set were high. Correlations between Dimension-1 and the dimension-1 of the three random sets were *r*(30) = .99, *p<.0001*; *r*(30) = .98, *p<.0001*; and *r*(30) = .99, *p<.0001*. Correlations between Dimension-2 and the dimension-2 of the three random sets were *r*(30) = .97, *p<.0001*; *r*(30) = .89 *p<.0001*; and *r*(30) = .92, *p<.0001*.

Because no reliable psychoacoustical measures could be derived for very short sounds, an additional MDS analysis on the grouping data was performed on the data excluding these sounds. This was done to be able to interpret the dimensions in terms of psychoacoustical measures. The MDS analysis again resulted in two dimensions with a Stress-I value of .11. These dimensions correlated high with Dimensions 1 and 2 resulting from the analysis containing all sounds, *r*(19) = .96, *p*<.0001 and *r*(19) = .95, *p*<.0001, respectively. The regression weights stemming from the PCA on the psycho-acoustical measures were correlated with the two dimensions from the MDS analysis that did not contain the short sounds.Dimension-1 correlated low with Factor 1 (*r*(19) = -.02, *p*=NS) and high with Factor 2 (*r*(19) =.87, *p*<.0001). Dimension-2 correlated higher with Factor 1 (*r*(19) = -.34, *p*=NS) than with Factor 2 (*r*(19) = -.11, *p* = NS). Thus, high values on *Dimension-1* are associated with 'low frequencies' and a higher level of noise (because Harmonicity loads negatively on Factor 1) and high values on *Dimension-2* are associated with higher levels of 'sharpness' and 'loudness'.

In Figure 1, the two dimensions of the MDS analysis are shown. Numbers in the figure indicate the product sounds presented in Table 1. As can be seen in this figure certain sounds seem to be grouped together that may reflect specific product sound categories. In order to determine these categories a hierarchical cluster analysis using Ward's method was conducted on the 2-dimensional coordinates. This yielded six relevant clusters (product sound categories). Each category contained at least four sounds. In Figure 1 the categories are indicated by density ellipses (*p*<.95) encircling the sounds.

Product sound category 1 contains short duration sounds caused by an impact between product parts (e.g., door closing). The sounds have a pulse-like character that on theoretical grounds (FFT) will result in a wide spectrum and also high frequencies. These sounds are positioned the lowest on Dimension 1 and relatively high on Dimension 2, which means they will evoke a sense of sharpness. The sounds were described with terms like: 'door, switch, short, single, click, bang, opening, closing'. Consequently, this category was named *Impact* sounds.

Product sound category 2 contains mostly digitally produced alarm-like sounds. In Figure 1, it can be seen that these sounds are positioned relatively low on Dimension 1, which means they are not noisy sounds, and are positioned relatively high on Dimension 2 which means they are loud and sharp sounds. The sounds were described with terms like: 'bell, beep, buttons on a microwave oven, warning, alarm, attention'. Because the majority of the descriptions indicated an alarming situation, this category was named *Alarm* sounds.

Product sound category 3 contains an old-fashioned alarm clock bell, a phonograph winding, a shaver, and an electric toothbrush sound. The products in this cluster are rather small in size. These sounds are the consequences of engines with high RPM, small rotating and rubbing mechanical parts of products. The sounds are positioned at the mid-point of Dimension 1, which means they have some noisiness in their spectral content, and are positioned relatively high on Dimension 2, which means they are sharp and loud sounds. The sounds were described with terms like: 'adjusting, rotating, rattling, shaver, buzzer, engine, machine, mechanism, mechanical'. Because the descriptions refer to mechanism related events and products that involve mechanical structures, this category was called *Mechanical* sounds.

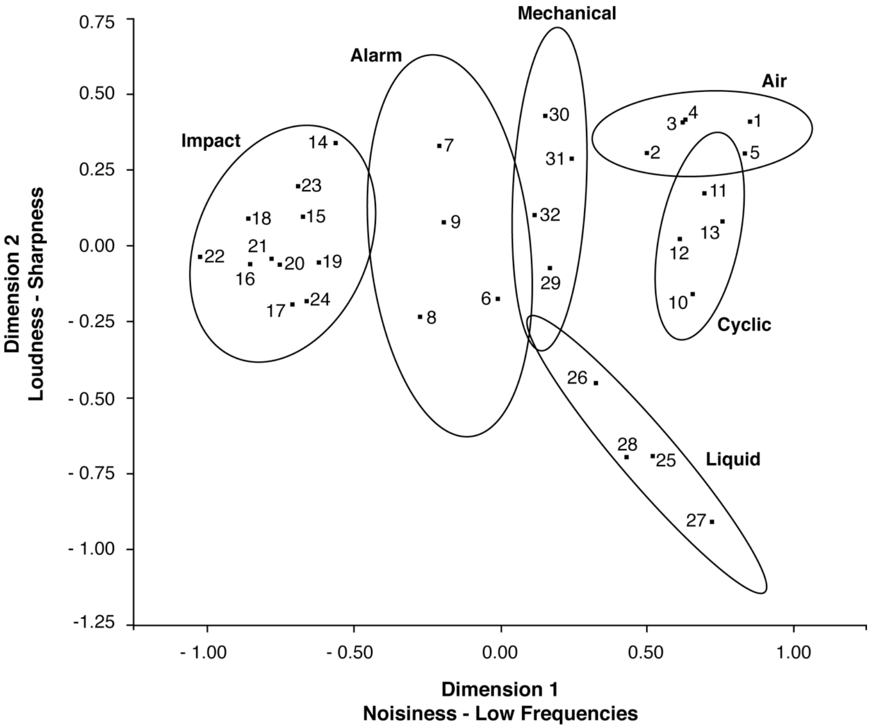


Figure 1. Product sound categories and their underlying dimensions. Dimension 1 represents Noisiness and Low-frequencies and Dimension 2 represents Loudness and Sharpness.

Product sound category 4 contains coffee machine sounds (coffee boiling, water heating, water pouring) and an electric kettle (water boiling) sound. The products in this category contain liquids. These sounds are mostly caused by the heating of liquids. The sounds are positioned relatively high on Dimension 1, which means they contain low-frequencies in their spectral content and are somewhat noisy, and are positioned the lowest on Dimension 2, which means they are rather quiet and not sharp sounds. The sounds were described with terms like: 'coffee machine, water, coffee, boiling, pouring, filling, bubble'. Because the descriptions indicate liquid related events, this cluster was named '*liquid'* sounds.

Product sound category 5 contains microwave oven, kitchen hood, dishwasher, and tumble-dryer sounds. These products employ rotating parts which cause a cyclic event and a periodicity in the sound as a result. Some of the products in this category employ liquids. The sounds in this category are positioned relatively high on Dimension 1, which means they are rather low and noisy sounds, and are positioned relatively low on Dimension 2 compared to category 6, which means they are also loud and sharp. The sounds were described with terms like: 'vacuuming, blowing, dryer, fan, monotone, soft, low pitch, laundry room, ventilator, background'. These descriptions resemble the descriptions of the sounds in Category 6, however these sounds can be distinguished from them by their fluctuating temporal properties which indicate a rotating event. Therefore, this category was named '*cyclic'* sounds.

Product sound category 6 contains vacuum cleaner, hand vacuum cleaner, washing machine, washing machine centrifuge cycle, and hair dryer sounds. These products are highly involved with air interaction due to the rotating fans used to blow or suck air. The sounds are positioned the highest on Dimension 1 which means they are noisy and consist of low frequencies and highest on Dimension 2 which means they are relatively sharp and loud sounds. The sounds were described with terms like: 'vacuum cleaner, hair dryer, air, drying, blowing, vacuuming, aerodynamic. Because all these descriptions indicate events and products related to air, this category was named '*air'* sounds.

Representativeness

To determine the agreement between the participants on the representativeness, entropy measures were obtained for each sound category using Shannon's index for diversity (Zar, 1996). The entropy measure for *Liquid* sounds was the lowest (.47). This indicates that listeners chose the same representative sound more often than the other sounds. The entropy measure for *Impact* sounds was the highest (.88), which indicates that there was a large dispersion in indicating the representative sound. Impact sounds were followed by *Air* sounds (.68). The entropy measures for *Cyclic* (.52), Alarm (.53), and Mechanical sounds (.54) were similar.

Participants chose 148 sounds as the representative member of a category. Sound 28 (coffee machine water pouring) was chosen 16 times (N = 148; n = 16), Sound 32 (phonograph winding) (N = 148; n = 14), Sound 6 (alarm clock beep) (N = 148; n = 13) are the sounds that were selected most frequently as a representative sound for liquid, mechanical, and alarm categories respectively. On the other hand, Sound 24 (sewing machine needle click) was never selected as a representative sound. Sound

11 (kitchen extractor fan) and sound 13 (tumble dryer) were selected seven times as a representative sound in cyclic sounds. Sound 1 (washing machine centrifuge cycle) was selected only three times as a representative sound in air sounds. Sound 19 (kitchen extractor fan power button) and sound 17 (microwave oven door opening) were selected as a representative sound in the impact sound category.

Category labels

Free labeling of the categories yielded various sound descriptions. To determine the basic concepts, these descriptions were classified. If a description consisted of more than one word, it was split up into meaningful sections (e.g., ‘unpleasant mechanical sounds’ as ‘unpleasant’ and ‘mechanical’). The resulting words were analyzed to determine conspicuous patterns in product sound descriptions. This analysis resulted into nine different concepts: action, emotion, location, meaning, onomatopoeia, psychoacoustics, sound type, source, and temporal.

Figure 2 presents the relative frequency of words as a function of basic concepts and product sound categories. In this figure, the bar indicated with ‘overall’ represents the relative frequency of words for all product sound categories. This overall measure shows that *source* descriptions and *onomatopoeias* are the most frequently used concepts (21.69%), followed by *action* (12.98%) and *sound type* (9.29%) concepts. (For the explanation of the concepts and the characterizing descriptions of the product sounds, see Özcan & van Egmond, 2005).

For the individual product sound categories, *air* sounds were mostly described by sound type, sound source, and location descriptions. Whereas, meaning descriptions were hardly used. *Alarm* sounds were mostly described by meaning, onomatopoeia, psycho-acoustics, and emotion descriptions. Whereas, action descriptions were hardly used sound descriptions and location descriptions never. *Cyclic* sounds were mostly described by sound source, sound type, psycho-acoustics, and location descriptions. Whereas, meaning descriptions were hardly used. *Impact* sounds were mostly described by onomatopoeia, action, sound source, and temporal descriptions. Whereas, location and the sound type descriptions were hardly used. *Liquid* sounds were mostly described by source, onomatopoeia, action, and emotion descriptions. Whereas, location descriptions were hardly used. *Mechanical* sounds were represented by multiple concepts such as by onomatopoeias, meaningful associations, sound source, sound type, and action descriptions. Whereas, location and psycho-acoustical descriptions were hardly used.

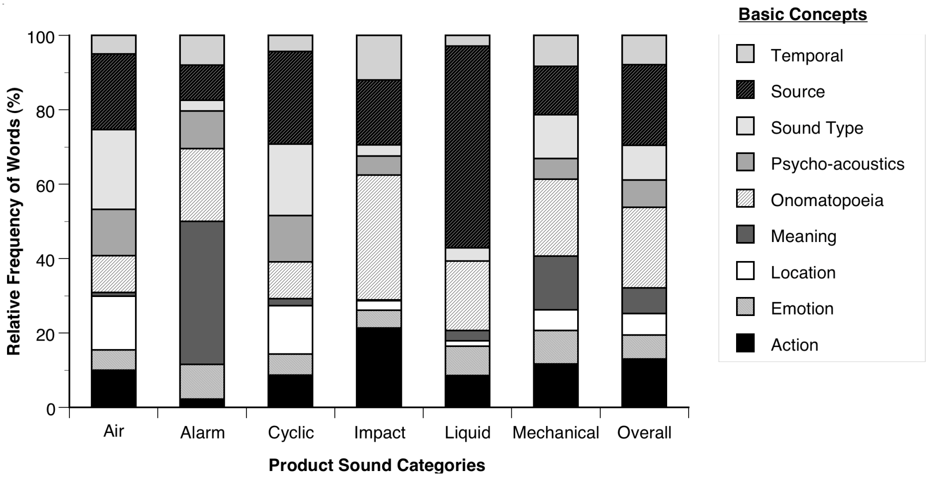


Figure 2. Relative frequency of words as a function of basic concepts and of product sound categories. (The cumulative percentage over descriptive groups add up to 100%).

Discussion

Six categories have emerged within the domain of product sounds. These product sound categories are identified as ‘air’, ‘alarm’, ‘cyclic’, ‘impact’, ‘liquid’, and ‘mechanical’. The occurring product sound categories resemble the categories proposed for the environmental sound domain such as air, liquid, impacting solids, complex events such as machines, and signaling sounds (Ballas, 1993; Gaver, 1993a; Gaver, 1993b). Interestingly, our sound categories relate both to Gaver’s theoretical account and Ballas’s empirical approach that employed cognitive judgments.

We have determined the (psycho)acoustical correlates of the product sound similarity to understand on what physical grounds the sounds differentiated from each other. We obtained two dimensions on which these categories are positioned. The first dimension is associated with *low frequencies* and *noisiness/harmonicity*, and the second dimension is positively associated with *sharpness* and *loudness*. Dimension 1 seems to distinguish between the categories. Categories are positioned on Dimension 1 in the following order: impact, alarm, mechanical, liquid, cyclic, and air. Whereas, Dimension 2 seems to distinguish between the category members, and also between air, liquid, and cyclic sounds.

The acoustical differentiation of the product sounds is not satisfactory to understand the semantic differentiation. The sound descriptive words produced during the labeling task provided us with some insight into mental representation of sounds. Nine different basic concepts were derived from those words: action, emotion, location, meaning, onomatopoeia, psychoacoustics, sound type, source, and temporal. It has been shown that product sound categories can be interpreted in terms of these basic concepts. *Source* descriptions and *onomatopoeias* appear to be the most important concepts. Thus, listeners try to identify a sound by its source. If that is not possible, mimicking the sound is the second strategy (i.e., onomatopoeias). In other words, listeners are unable to extract enough semantic association about the perceived sound, suggesting that product sounds are not always well identified. Alarm, impact and liquid sounds are the categories that are most clearly related to a specific concept. Liquid sounds are most related to a *source*. Impact sounds are related to *onomatopoeias*, because they lack structure and are very short, thus, do not convey much information. Impact sounds also elicit *action* and *temporal* descriptions. Alarm sounds are mostly related to a *meaning* concept. This is understandable because these sounds are designed to convey a message. Air and Cyclic sounds are differentiated better with *location* descriptions. The remaining categories are related to more than one concept. In conclusion, each product sound category is acoustically and semantically distinguishable. Moreover, some categories are represented better with certain concepts.

These mental representations also reveal how product sounds are in general processed without being affected by external conditions (e.g., context). The processing of a product sound may (a) stop at the perceptual analysis phase or (b) result in a cognitive judgment. Perceptual analysis results in sensorial judgments related to the psychoacoustical properties of a sound and in a structural analysis that produces temporal descriptions and onomatopoeias. Cognitive processing results in information about the source and the action causing the sound and the location where it can be heard. Meaningful associations and sound type are also a result of cognitive processing. Furthermore, these perceptual or cognitive processing may result in emotional response to the perceived sound.

In Experiment 1, a wide range of product sounds has been used. Categorization of product sounds may have resulted in clear and distinct categories of sounds. Experiment 2 will investigate the categorization criteria for perceptually similar sounds.

Experiment 2

Categorization of perceptually similar product sounds

The *air*, *cyclic*, and *mechanical* sound categories are more similar to each other than alarm, impact, and liquid sounds (see Figure 1). The similarity between *air* and *cyclic* sounds is probably caused by the lower frequencies in the spectral content, but they are dissimilar because of the differences in loudness and sharpness. The *mechanical* sounds are less noisy and consist of higher frequencies compared to air and cyclic sounds. Moreover, Experiment 1 suggests some overlapping category memberships. For example, the washing machine sound (Sound 5 in Figure 1) seems to belong to the *air* sound category, but it is also placed close the category border of the *cyclic* sounds. This suggests that due to the high perceptual similarity, some sounds may belong to more than one category. Multiple category memberships may be indicative of ambiguity in sound identification. Therefore, we predict that high perceptual similarity between product sounds will hinder the correct identification of the sounds and listeners will rely more on, e.g., the psycho-acoustical judgments in order to distinguish between them (see Vanderveer, 1979) and less on the sound source. That is because two acoustically very similar sounds will evoke similar conceptual associations and probably the same lexical association. Thus distinguishing between these sounds by source information may not be a strategy a listener will use. However, a listener may use subtle spectral-temporal changes in the sound to determine how the sound will be categorized.

Therefore, Experiment 2 investigates listeners' ability to distinguish between perceptually similar sounds and the dimensions on which listeners base their judgment. The procedure was identical to Experiment 1. Thirty-two product sounds were selected to represent three of the six product sound categories (i.e., air, cyclic, and mechanical) defined in Experiment 1. In addition to the psychoacoustical metrics, six attributes (sharp, rough, hard, dull, calm, regular) were rated in order to determine on which perceptual aspects sounds are categorized.

Method

Participants

Eighteen students from Delft University of Technology volunteered in the categorization task. Sixteen different students participated in the subjective rating task. The average age was 21.6 years. All participants reported normal hearing.

Stimuli

Thirty-two product sounds were selected to represent three sound categories defined in Experiment 1 (air, cyclic, and mechanical sounds). In order to create high perceptual similarities between the sounds, more than one recording of a product type was included in the stimulus set (e.g., the sound of three different juicers was recorded). The sounds were recordings of various electrical domestic appliances in operation. Only four sounds were selected from Experiment 1; other 28 sounds were recorded in house conditions by using a recording apparatus, Boss BR-532, with a

Sennheiser e865 microphone with a frequency response of 40Hz - 20kHz and free-field sensitivity of 3mV/Pa. The sounds were edited on a Macintosh PowerPC G4 computer using the sound-editing program Sound Studio, sounds longer than 5 seconds were trimmed to a maximum duration of 5 seconds. All sounds were saved in a stereo format with a sampling rate of 44.1 kHz and 16 bits. The loudness levels were adjusted to a comfortable listening level for each sound. The participants were not allowed to change the sound levels during the experiment.

Table 2 presents the psycho-acoustical parameters measured and calculated for each product sound. The sound pressure level (SPL) of each sound was measured by a Bruel & Kjaer 2260 Investigator. The SPL level was used to calibrate the psycho-acoustical analysis software. Harmonicity was calculated using Praat[1]. Two sharpness parameters (Zwicker & Fastl, 1990; Aures, 1985), the spectral centroid, intensity, and 39 critical band levels in Erbs were calculated using Psysound[2]. To reduce the number of parameters a principal component analysis (PCA) with Varimax rotation was conducted that resulted in two factors explaining 69% of the variance. The two sharpness parameters, the spectral centroid, and high frequency critical bands (CB-3544 through CB-15085) high and positively, whereas, low frequency critical bands (CB-55 through CB-44) high and negatively loaded on *Factor 1* (explaining 43.35% of the variance). Therefore, this factor was interpreted as sharpness. The SPL, intensity, and mid-range frequencies (CB-520 through CB-3158) high and positively loaded on *Factor 2*, whereas Harmonicity loaded low and negatively on *Factor 2* (explaining 25.78 of the variance). Therefore, this factor was interpreted as loudness. In addition, the regression weights for *Factor 1 and Factor 2* were extracted for each sound.

Thirty-two sounds were rated on a 7-point scale for the six attributes (rough, sharp, dull, hard (related to the loudness), quiet, regular (referring to the temporal aspects)). The mean ratings of the attributes for each sound are presented in Table 3.

Apparatus and procedure

The stimuli were presented using the same software application from Experiment 1. The application ran on a Macintosh Power PC G4 computer using MacOSX with a 17” Iiyama Pro454 monitor. Apple Pro Speakers with a frequency range of 70Hz to 20kHz and a maximum power of 10 Watts per channel were used to monitor the stimuli. The study took place in a quiet room. The procedure was identical to Experiment 1.

			Psychoacoustical metrics					
Cluster	**ID**	**Sound Description**	**SPL**	**SC**	**S(Z&F)**	**S(A)**	**Int. (E-05)**	**Harm.**
1	1	citrus press	76	750	1.2	3.0	3.98	1.05
	2	citrus press	75	735	1.3	3.3	3.16	-0.11
	3	can opener	71	1894	1.5	3.8	1.26	-4.41
	4	citrus press	74	1005	1.3	2.9	2.51	9.44
2	5	mixer	75	1736	1.6	4.6	3.16	-2.03
	6	mixer	74	2270	1.7	4.6	2.51	3.85
	7	hairdryer	75	1621	1.6	4.2	3.16	-0.96
	8	vacuum cleaner (hand)	75	1671	1.6	4.2	3.16	-1.76
3	9	computer	71	485	1.2	2.9	1.26	2.17
	10	fridge	69	119	1.0	1.8	0.79	14.64
	11	microwave	74	233	1.0	1.9	2.51	2.67
	12	tumble dryer	73	117	0.9	1.5	2.00	6.93
	13	kitchen extractor fan	77	390	1.5	3.3	5.01	1.33
	14	washing machine	73	173	1.0	1.9	2.00	2.47
	15	washing machine	72	258	1.3	2.9	1.58	3.20
	16	tumble dryer	73	204	1.0	2.0	2.00	0.38
	17	dishwasher	73	272	1.3	2.9	2.00	-1.57
4	18	hand blender	72	1955	1.6	4.3	1.58	13.31
	19	shaving machine	74	3972	2.0	5.6	2.51	-0.58
	20	shaving machine	72	2095	1.7	4.6	1.58	17.42
	21	hair clippers	72	2748	1.8	4.9	1.58	21.67
5	22	hand blender	72	1131	1.5	3.7	1.58	9.77
	23	hair clippers	74	242	1.5	3.5	2.51	23.05
	24	toothbrush	68	3260	1.9	4.8	0.63	5.55
	25	toothbrush	72	2914	1.8	4.8	1.58	0.85
	26	toothbrush	71	3501	1.9	5.0	1.26	-1.71
	27	pedicure machine	70	4165	2.0	5.2	1.00	-2.37
6	28	vacuum cleaner	76	1326	1.5	4.4	3.98	-5.13
	29	hairdryer	77	1853	1.7	4.7	5.01	0.07
	30	centrifuge cycle (WM)	73	638	1.5	3.7	2.00	5.49
	31	vacuum cleaner (hand)	72	1284	1.5	3.7	1.58	4.94
	32	washing machine	71	1673	1.7	4.4	1.26	-2.05

Note. ID = sound number used in Figure 1, SPL = sound pressure level measured in decibels, SC = spectral centroid, S(Z&F) = sharpness algorithm defined by Zwicker & Fastl (1993), S(A) = sharpness algorithm defined by Aures (1985), Int. = sound intensity (in W/m^2), Harm. = harmonicity. WM = washing machine.

Table 2. Psychoacoustical metrics calculated for similar product sounds.

Results

The minimum number of categories created was three and the maximum number of categories was nine in the grouping task. Six participants created five categories. The mean category fit was 4.38 and differed significantly from the middle-point of the scale (*t(17)=2.68, p<.01*). This indicated that on average the participants were satisfied with their groups. The method of analysis was identical to that of Experiment1.

A multidimensional scaling analysis yielded a 2-dimensional solution with a Stress-I value of .14. The 3-dimensional solution resulted in a minor decrease in stress (from .14 to .11), and was therefore not used. In addition, the effect of sample size on the solution was tested. Three randomly selected samples (n=9) were drawn from total dataset (N=18) and analyzed using the aforementioned MDS procedure. The correlations between the coordinates of these three sets and the coordinates of the entire set were high. Correlations between Dimension-1 and three random sets were r(30) = .95, *p<.0001*; r(30) = .92, *p<.0001*; and r(30) = .99, *p<.0001*. Correlations between Dimension-2 and three random sets were r(30) = .91, *p<.0001*; r(30) = .76 *p<.0001*; and r(30) = .83, *p<.0001*. Thus, the sample size was sufficiently large to obtain a stable solution.

A stepwise regression analysis was conducted to interpret the dimensions in terms of the psychoacoustical metrics and subjective ratings. The regression weights resulting from the Principal Component Analysis and subjective ratings were used as independent variables to predict Dimension 1 and Dimension 2. The enter criteria was $p < .05$.

For Dimension 1, the results showed that the total amount of variance explained was 40% with the subjective ratings Calm and Rough [F(2, 29) = 10.07, $p < .001$] of which 15% [given by the semipartial coefficient of determination] was attributable to Calm (t(28) = - 2.72, p<.05) and 39% [given by the semipartial coefficient of determination] was attributable to Rough (t(28) = - 4.37, p<.001). The estimated regression function for Dimension 1 (using standardized beta coefficient) was:

$$\text{Dimension 1} = 4.48 - 0.61\ \text{Calm} - 0.98\ \text{Rough}$$

Thus, an increase in calm or in rough results in a low positioning on Dimension 1.

For Dimension 2, the results showed that the total amount of variance explained was 77% with the subjective rating Hard and Loud (Factor 2) [F(2, 29) = 47.52, $p < .001$] of which 39% [given by the semipartial coefficient of determination] was attributable to Hard (t(28) = 3.85, $p < .05$) and 12% [given by the semipartial coefficient of

			Attributes					
Cluster	**ID**	**Sound Description**	**Sharp**	**Rough**	**Hard**	**Quiet**	**Dull**	**Regular**
1	1	citrus press	4.35	5.35	5.29	3.12	4.18	4.00
	2	citrus press	4.35	5.06	4.82	3.24	4.06	4.06
	3	can opener	4.35	5.00	4.35	3.29	3.88	3.71
	4	citrus press	4.94	5.12	5.00	2.76	4.12	3.47
		Mean	*4.50*	*5.13*	*4.87*	*3.10*	*4.06*	*3.81*
2	5	mixer	4.82	4.24	4.76	2.59	4.47	4.24
	6	mixer	5.82	4.47	5.53	2.41	4.65	4.53
	7	hairdryer	5.06	3.94	4.82	3.24	4.59	4.24
	8	vacuum cleaner (hand)	5.41	4.00	4.76	2.94	4.12	3.59
		Mean	*5.28*	*4.16*	*4.97*	*2.79*	*4.46*	*4.15*
3	9	computer	3.35	3.71	3.06	4.47	4.53	4.29
	10	fridge	1.76	3.00	1.94	5.18	4.65	4.88
	11	microwave	1.76	2.35	1.71	5.00	4.18	4.59
	12	tumble dryer	2.35	3.24	2.53	4.59	4.47	4.06
	13	kitchen extractor fan	2.29	3.24	2.82	4.59	3.88	4.35
	14	washing machine	2.35	3.53	2.59	4.29	3.88	4.00
	15	washing machine	2.82	3.29	2.47	4.47	4.41	4.12
	16	tumble dryer	2.29	3.00	2.29	4.94	4.47	4.47
	17	dishwasher	2.29	3.41	2.47	3.71	3.00	3.24
		Mean	*2.36*	*3.20*	*2.43*	*4.58*	*4.16*	*4.22*
4	18	hand blender	6.12	4.65	5.35	2.29	4.53	4.47
	19	shaving machine	5.76	4.59	4.76	3.00	4.59	4.12
	20	shaving machine	5.71	5.41	5.88	2.06	4.06	3.94
	21	hair clippers	5.88	4.94	5.88	2.65	4.06	4.53
		Mean	*5.87*	*4.90*	*5.47*	*2.50*	*4.31*	*4.26*
5	22	hand blender	4.12	4.12	3.76	3.47	4.47	4.53
	23	hair clippers	4.06	4.82	4.53	3.53	4.65	5.41
	24	toothbrush	4.88	4.94	4.88	3.00	4.29	4.53
	25	toothbrush	4.53	4.18	4.53	3.12	4.35	4.24
	26	toothbrush	5.00	4.82	4.59	3.94	4.24	3.65
	27	pedicure machine	4.41	4.65	3.94	3.76	4.12	3.65
		Mean	*4.50*	*4.59*	*4.37*	*3.47*	*4.35*	*4.33*
6	28	vacuum cleaner	4.35	4.53	3.88	3.35	4.35	4.88
	29	hairdryer	4.59	4.71	5.06	3.35	4.53	4.65
	30	centrifuge cycle (WM)	5.06	4.00	4.24	2.53	4.00	2.82
	31	vacuum cleaner (hand)	4.88	4.47	4.41	3.06	4.76	4.88
	32	washing machine	4.76	4.41	4.12	3.47	4.18	3.76
		Mean	*4.73*	*4.42*	*4.34*	*3.15*	*4.36*	*4.20*

Note. ID = sound number used in Figure 1, WM = washing machine.

Table 3. Subjective evaluations of similar product sounds

determination] was attributable to Loud ($t(28) = 6.97$, $p < .001$). The estimated regression function for Dimension 2 (using standardized beta coefficient) was:

$$\text{Dimension 2} = -.93 + 0.37\ \text{Loud} + 0.67\ \text{Hard}$$

Thus, an increase in Loud and Hard results in higher positioning on Dimension 2.

In Figure 3, the two dimensions of the MDS analysis are shown. Numbers in the figure indicate the product sounds. As can be seen in this figure certain sounds seem to be grouped together that may reflect specific subgroups within a product sound category. In order to determine these sub-groups a hierarchical cluster analysis using Ward's method was conducted on the 2-dimensional coordinates. This yielded six relevant clusters (sub-*sound*-groups). Each category contained at least four sounds. In Figure 3 the categories are indicated by density ellipses ($p<.95$) encircling the sounds.

Sound group 1 contains mainly three citrus press sounds and a can opener sound. These sounds are a sub-category to mechanical sounds and characterized by high roughness ratings. *Sound group 2* contains two mixer sounds, plus a hairdryer and a vacuum cleaner sound. These sounds are a sub-category to air sounds and characterized by high sharpness and hardness, and low calmness ratings. *Sound group 3* contains sounds of large size products (i.e., two washing machine, two tumble dryer, microwave oven, dishwasher, fridge, computer, and kitchen extractor fan sounds). These sounds are a sub-category to cyclic and air sounds and characterized by the highest calmness and the lowest sharpness, roughness, and hardness ratings. *Sound group 4* contains two shaver sounds, a hair-clipper and a hand blender sound. These sounds are a sub-category to mechanical sounds and characterized by the highest sharpness and hardness, and the lowest calmness ratings. *Sound group 5* contains three toothbrush sounds, hair-clippers, a pedicure machine, and a hand blender sound. These sounds are a sub-category to mechanical sounds and characterized by the highest regularity ratings. *Sound group 6* contains two vacuum cleaner and washing machine sounds, and a hairdryer sound. These sounds are a sub-category to air sounds and characterized by high sharpness and low calmness ratings.

Sound descriptions

The category labels were analyzed the same way as in Experiment 1. This analysis resulted in ten different categories of product sound descriptions: action, emotion, location, meaning, onomatopoeia, psychoacoustics, sound type, source, source properties, and temporal descriptions. The relative frequency of words as a function of basic concepts for Experiment 2 is presented in Figure 4. According to the figure,

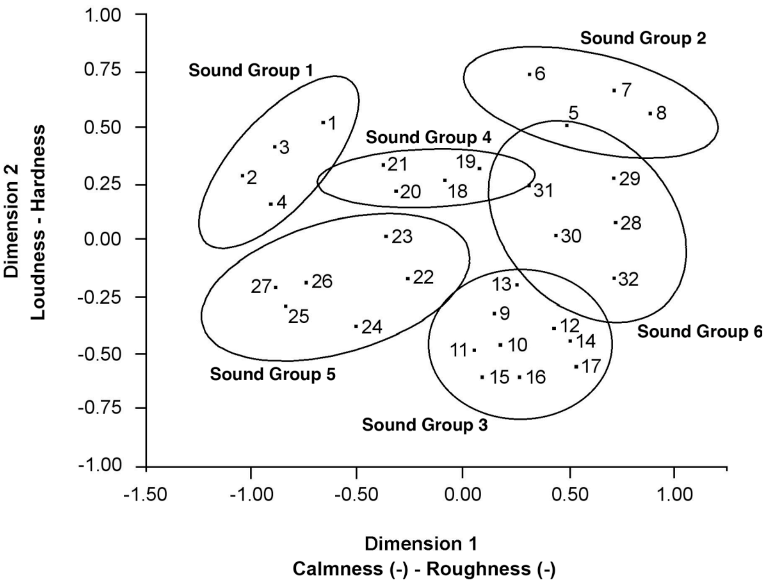


Figure 3. Underlying dimensions of similar product sounds. Dimension 1 represents Calmness (-) and Roughness (-) and Dimension 2 represents Loudness and Hardness.

source descriptions are the most frequently used sound description category with 34.45% of all the descriptions, followed by *location* (15.87%) and *psycho-acoustics* (13.66%) descriptions. The least frequently used descriptions are *emotion* (2.58%)descriptions. A new type of sound description has emerged in Experiment 2: *source properties*. Source properties constituted 6.37% of the sound descriptive words.

To be able to compare the sound descriptive words with Experiment 1, the relative frequency of descriptive sounds were calculated for air, cyclic, and mechanical sounds (see Figure 4). In Figure 4, the frequency of sound source descriptions is higher (by 14.72%) for Experiment 2; and source property descriptions only occurred in Experiment 2. The frequency of action descriptions remained approximately the same for Experiment 2. The frequency of sound type descriptions noticeably decreased (10.81%) in Experiment 2. The choice of sound descriptive words significantly differed between Experiment 1 and Experiment 2 (X^2 (8, N=1166) = 118.11, p<.0001).

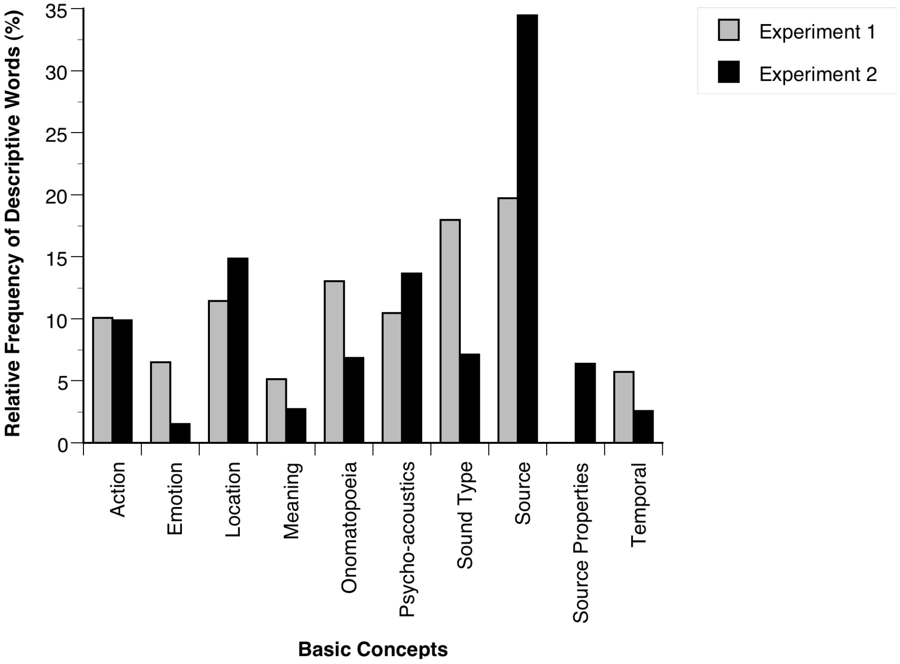


Figure 4. Comparison of Basic Concepts from Experiment 1 and Experiment 2.

Discussion

Experiment 2 has been mainly conducted to determine the underlying perceptual dimensions in the categorization of very similar product sounds. Results indicate that four main psycho-acoustical measures underlie the categories of perceptually similar sounds. These are Roughness and Calmness for Dimension 1 and Loudness and Hardness for Dimension 2.

For Dimension 1, sounds that have been rated rough and calm are mainly the sub-groups of mechanical sounds (Groups 1, 4, and 5); the sub-groups of air and cyclic sounds (Groups 2, 3, and 6) have not been rated that rough. The calmness attribute exhibits a similar trend. However, the results of the regression analysis indicate that it is mainly the roughness attribute that characterizes Dimension 1. For Dimension 2, sounds from Group 1 (mechanical sounds) are relatively high in roughness and hardness, and relatively low in calmness; sounds from Group 2 (air sounds) are relatively high in sharpness and hardness, and relatively low in calmness. Sounds from Group 3 (air and cyclic sounds) are the lowest in sharpness, roughness, and hardness and the highest in calmness. The results of the regression analysis indicate that the hardness attribute characterizes Dimension 2 the most. Although not conclusive, we can speculate that high sharpness and roughness ratings may somewhat be indicative for high hardness ratings.

The sounds have been selected to represent each of the three product sound categories (air, cyclic, and mechanical). The six groups of Experiment 2 are better differentiated than the three sound categories in Experiment 1. These groups can be considered as the sub-groups of these product sound categories. The difference between the sub-groups can be described by the underlying psycho-acoustical properties. These indicate that psycho-acoustical similarity dominates the sound categorization for similar sounds, which has also been found by Vanderveer (1979). In Experiment 1 the sound categories could be distinguished by the auditory properties capturing the interaction of materials (solid, liquid, air, digital).

Listeners tend to associate the perceptually similar product sounds more frequently with their sources in Experiment 2 than in Experiment 1. This probably stems from the multiple recordings of the same product type (e.g., three recordings of different toothbrushes). An alternative approach may suggest that similar sounds activate the same lexical association (i.e., source information). Consequently, such sounds could have been differentiated by psychoacoustical judgments (see the increase for psychoacoustical judgments in Figure 4), but identified by the same sound source. That is because, perceptual processes result on a more reliable structural distinction between the sounds, as cognitive processes lead to the same lexical representation. Thus, an active featural comparison may have been a better strategy. Therefore, two types of judgment strategies have emerged: holistic approach (i.e., naming the source) and a feature analysis approach (i.e., differentiation by the psychoacoustical attributes). In the next experiment, we will investigate more thoroughly which verbal attributes relate to product sounds.

Experiment 3

Verbal attributes of product sounds

Experiments 1 and 2 have shown that when perceiving a product sound a semantic store in memory is accessed. In order to explore the semantic associations that are activated when perceiving a sound, listeners were asked to rate multiple attributes for the same sound in Experiment 3. By doing this, we will also determine the extent to which certain labels characterize a product sound.

Listeners were asked to rate 18 domestic product sounds on 48 semantic attributes. The attributes were selected from other studies that used environmental, synthetic, musical, and machinery sounds (Edworthy et al., 1995; Kendall & Carterette, 1993; Solomon, 1958; von Bismarck, 1974), or from the sound descriptions provided by participants in Experiments 1 and 2. (These attributes are categorized into nine groups: action, location, material, meaning, psychoacoustics, sound type, source, source properties, and temporal descriptions).

Method

Participants

Thirty-six students of the Delft University of Technology participated. The average age of the participants was 23.3 years. All participants reported normal hearing. Students voluntarily participated and were paid.

Stimuli

Three sounds from Experiment 1 were selected from each of the six perceptual product sound categories. Thus, in total 18 product sounds were selected. The sounds had a maximum duration of 5 seconds and were presented at a similar comfortable listening level.

Rating attributes

Forty-eight attributes (excluding the italic ones) are presented in Table 4 as a function of basic concepts (column names). In the table the marks next to the words refer to the studies in which they were previously used. The words of the ‘psychoacoustics’ category partly reflect the Roughness / Sharpness and Loudness dimensions from Experiment 2. The material descriptions are included because Experiment 2 indicated that product sound categories are described by the material interaction of the sources. Source descriptions are not provided; instead two attributes that refer to participants’ familiarity with the sources are chosen. The words were translated from English into Dutch.

Apparatus

The stimuli and the descriptive words (attributes) were presented using a specially designed application developed using the Trolltech Qt (Mac OS X - free edition) tool kit. The application ran on a Macintosh Powerbook G4 1.33 GHz computer with 12" screen. The stimuli were presented through AKG Studio Monitor K240DF 2x600Ohm headphones. The experiment took place in a quiet room.

Procedure

A participant’s task was to rate the presented product sound for each attribute on a 7-point unipolar scale (‘1’ representing “weak” and ‘7’ representing “strong” association) or a participant could indicate “non- applicable’ (N/A). The latter choice indicated that there was absolutely no semantic association between the attribute and the

Action	Emotion	Location	Material	Psycho-acoustics	Sound Type	Source	Source Properties	Temporal Descriptions
Opening	*Tense* * † ^	Kitchen	Metal	Sharp * † ^	Digital	Familiar †	Big	Short
Closing	*Relaxed* * †	Bathroom	Plastic	Dull * †	Electrical	Strange †	Small †	Long
Pouring	*Obtrusive* *	Bedroom	Glass	High-pitched	Mechanical		Massive	Repetitive
Blowing	*Reserved* *		Wood	Low-pitched	*Machine*		Hallow * † ^	Constant
Droning	*Unpleasant* *		Rubber	Round * ^	*Manual*		*Weak* * † ° ^	Irregular
Rotating	*Pleasant* * † °		Hard	Rough * †			*Powerful* † °	Fast † °
Impacting			Soft	Pure * ^				Slow † °
			Solid * †	Noisy				
			Liquid	Loud * † ^				
			Air	Quiet				
				Edged				
				Smooth * † ° ^				

Note. The underscored words were used only in Experiment 3. The words in italics were used only in Experiment 4. The remaining words were common both in Experiment 3 and Experiment 4.

* Words occurred in von Bismarck's study (1974).

† Words occurred in Solomon's study (1958).

° Words occurred in Edworthy et al.' s study (1995).

^ Words occurred in Kendall and Carterette's study (1995).

Table 4. Selected attributes and product sound related basic concepts.

corresponding sound. The instructions were presented on an A4 paper prior to the experiment. Age, gender, and nationality of a participant were collected on a small questionnaire. First, a participant took a practice trial with two animal sounds and three attributes that were not used in the study. Participants were encouraged to listen to each sound more than once during the rating task. Following the practice trial, a participant started the real experimental session. The presentation order of the stimuli and of the attributes was randomized for each participant.

Results

Data were analyzed in three phases: elimination of non-applicable rating attributes, factor analysis, and reliability tests on the descriptive word of each extracted factor.

Elimination of attributes

Elimination of the attributes was performed in three stages. First, participants that often rated ‘N/A’ were determined. On average, in 10% of the cases, a participant rated ‘N/A’. Two participants with ‘N/A’ rating frequencies of 48% and 57% were excluded from further analysis.

Second, for each product sound category, the words that were associated most frequently with N/A ratings were determined. A correspondence analysis on the frequency data of word-rating combination was performed and followed by a hierarchical cluster analysis (Ward’s method) on the coordinates stemming from the correspondence analysis. For four out of the six sound categories (air, cyclic, impact, and mechanical), cluster analysis yielded a cluster of words that were associated with N/A and ‘1’. For the other two sound categories (alarm and liquid), cluster analysis yielded a cluster of words that were associated with N/A only. In Table 5, the words as a function of product sound category are presented.

Third, it was determined if the attributes differentiated over sounds and if participants agreed. A measure of dispersion for the sounds was determined by taking the mean over participants for each attribute and sound, and then, the variance was calculated over sounds for each attribute. This variance indicated if attributes differentiated between sounds. A measure of dispersion was determined by taking the mean over sounds for each attribute and each participant, and then, the variance was calculated over participants for each attribute. This variance indicated how well participants agreed. Thus, each word (i.e., attribute) was associated with two variances. A hierarchical cluster analysis (Ward’s method, on the standardized data) was conducted with the two variances as input. Resulting cluster contained six words having a low agreement and a high differentiation measure. Three of these descriptive words had already been determined as an inappropriate word in the second stage. Thus, only three words were eliminated in this phase. Table 5 also

Analysis Type								***All Words***
Correspondence							**Variance**	
Air	**Alarm***	**Cyclic**	**Impact**	**Liquid***	**Mechanical**	***All***		
-	Air	-	-	-	Air	*Air*	-	*Air*
-	Bathroom	-	-	-	-	*Bathroom*	-	*Bathroom*
-	-	Bedroom	-	Bedroom	-	*Bedroom*	-	*Bedroom*
-	Blowing	-	Blowing	-	Blowing	*Blowing*	-	*Blowing*
Closing	-	Closing	-	Closing	Closing	*Closing*	Closing	*Closing*
Digital	-	Digital	-	Digital	Digital	*Digital*	-	*Digital*
-	Droning	-	Droning	-	-	*Droning*	-	*Droning*
-	-	-	-	-	-	-	Dull	*Dull*
Edged	-	Edged	-	-	-	*Edged*	-	*Edged*
Glass	Glass	Glass	Glass	-	Glass	*Glass*	-	*Glass*
Impacting	-	-	-	-	-	*Impacting*	-	*Impacting*
	-	-	-	-	-	-	Irregular	*Irregular*
-	-	-	-	-	-	-	Kitchen	*Kitchen*
Liquid	Liquid	-	Liquid	-	Liquid	*Liquid*	-	*Liquid*
-	-	-	Long	-	-	*Long*	-	*Long*
-	-	-	-	Metal	-	*Metal*	-	*Metal*
-	-	-	Noisy	-	-	*Noisy*	-	*Noisy*
Opening	-	Opening	-	-	Opening	*Opening*	Opening	*Opening*
-	-	Plastic	-	-	-	*Plastic*	Plastic	*Plastic*
Pouring	Pouring	-	Pouring	-	Pouring	*Pouring*	-	*Pouring*
Pure	-	-	-	-	Pure	*Pure*	-	*Pure*
-	Rotating	-	-	-	-	*Rotating*	-	*Rotating*
Rubber	Rubber	Rubber	Rubber	Rubber	Rubber	*Rubber*	-	*Rubber*
Short	-	Short	-	-	-	*Short*	-	*Short*
-	-	Small	-	-	-	*Small*	-	*Small*
-	-	-	Soft	-	Soft	*Soft*	-	*Soft*
Wood	Wood	Wood	-	Wood	Wood	*Wood*	-	*Wood*
Total number of words								
12	10	11	9	6	12	*24*	6	*27*

* The words in these categories were clustered only with 'N/A' ratings (not together with '1' ratings)

Table 5. Eliminated attributes in Experiment 3 for each product sound category in each analysis

presents the eliminated words as a function of analysis methods by which the inappropriate words were determined (i.e., correspondence analysis and the variance analysis). The last column (All Words) indicates the total number of words (27) that were excluded for further analysis. In the table, the word ‘rubber’ was never found appropriate to describe any sound category. The words ‘impacting’ was found inappropriate for only air sounds; ‘bedroom’ and ‘rotating’ for only alarm sounds; ‘small’ and ‘plastic’ for only cyclic sounds; ‘long’ and ‘noisy’ for only impact sounds; and ‘metal’ for only liquid sounds.

Factor Analysis

The ratings were analyzed using the method of principal components analysis with Varimax rotation. For the analysis, ‘N/A’ ratings were replaced by the mean values taken over participants and sounds for each attribute. Five factors explained 55% of the variance. A reliability analysis using Cronbach’s alpha model was conducted in order to check the internal consistency of the descriptive words in a factor. The Cronbach’s alpha values of each factor are presented in Table 6. These values ranged from .79 to .43. Factor 5 had a high alpha value (.73). The attributes loaded on this factor also had very high communalities (familiar: .87, strange: -.85) indicating a strong consistency within the factor. The factors, the attributes, and the explained variance are presented in Table 6. The factors can be interpreted as follows:

On *Factor 1,* words ‘low-pitch, slow, big, and quiet’ positively loaded high and words ‘high-pitch, sharp, and fast’ negatively loaded high. These words describe product sounds that stay quiet and rather unnoticed in the background. Therefore, Factor 1 was interpreted as *Inconspicuousness.* On *Factor 2*, words ‘hard, massive, mechanic, rough, and loud’ loaded high. Words describing the characteristics of a machine and the sound of it loaded on this factor. Therefore, Factor 2 was interpreted as the *Solidness* of the sound source. On *Factor 3*, words ‘repetitive, electrical, and constant’ positively loaded high. Factor 3 was interpreted as the *Repetitiveness*. On *Factor 4* words ‘smooth, round, and hollow’ positively loaded high. The words ‘smooth and round’ indicate the auditory quality of a sound and ‘hollow’ indicates a source property. Therefore, Factor 4 was interpreted as the *Smoothness*. On *Factor 5,* word ‘familiar’ positively loaded and word ‘strange’ negatively loaded high. These words indicate the familiarity of listeners with sound sources. Therefore, Factor 5 was interpreted as *Familiarity*.

The mean of the regression weights of the factor scores was determined over participants for each sound category. In Figures 5 A, B, and C, the sound categories are presented in three spaces for the 5 averaged regression weights. (In order to prevent abundance of data presentation, we have chosen to present only three combinations of factor dimensions). According to the figures, *air* sounds are

Attributes	Factor loading				
	1	2	3	4	5
Low-pitch	**0.71**	0.19	-0.09	0.04	-0.08
High-pitch	**-0.70**	0.16	0.26	0.09	0.19
Slow	**0.66**	-0.01	0.26	0.19	0.01
Fast	**-0.63**	0.26	-0.14	0.07	0.02
Sharp	**-0.62**	0.43	0.17	-0.07	0.04
Big	**0.61**	0.36	0.12	0.18	-0.01
Quiet	**0.51**	-0.35	-0.08	0.39	0.05
Hard	-0.29	**0.68**	-0.17	0.10	-0.03
Massive	0.17	**0.61**	-0.07	0.26	-0.10
Mechanic	-0.10	**0.59**	0.12	-0.02	-0.12
Rough	0.14	**0.58**	0.27	-0.26	0.03
Loud	-0.36	**0.53**	0.25	-0.01	0.27
Repetitive	-0.04	0.09	**0.74**	0.03	0.13
Electrical	-0.11	0.18	**0.66**	-0.24	-0.07
Constant	0.13	-0.09	**0.64**	0.17	0.01
Smooth	-0.19	-0.05	-0.02	**0.77**	-0.12
Round	0.23	0.05	0.13	**0.63**	0.07
Hollow	0.24	0.30	-0.10	**0.44**	0.09
Familiar	-0.08	0.03	0.09	0.03	**0.87**
Strange	0.08	0.15	0.00	0.03	**-0.84**
% of variance	16.55	12.70	9.13	8.36	8.35
α	0.79	0.63	0.55	0.43	0.73

Note. Boldface indicates highest factor loadings > .04. Factor 1 = Inconspicuousness, Factor 2 = Solidness, Factor 3 = Repetitiveness, Factor 4 = Smoothness, Factor 5 = Familiarity.

Table 6. Attributes, Factor Loadings for a Five-Factor Solution, Percentages of Variance Explained, and Cronbach's Alpha for the Attributes of Experiment 3.

positioned the highest on the Repetitiveness factor and rather high on the Solidness factor. *Alarm* sounds are positioned the highest on the Inconspicuousness and the Familiarity factors, are relatively high on the Repetitiveness factor. *Cyclic* sounds are positioned the lowest on the Inconspicuousness factor and relatively high on the Smoothness factor. *Impact* sounds are positioned the highest on Solidness factor and the lowest on the Repetitiveness and the Familiarity factor. *Liquid* sounds are positioned the highest on the Smoothness and the lowest on the Solidness factors. In addition, they are positioned relatively low on the Inconspicuousness and high on the Familiarity factors. *Mechanical* sounds are positioned relatively high on the Solidness and on the §Repetitiveness factors.

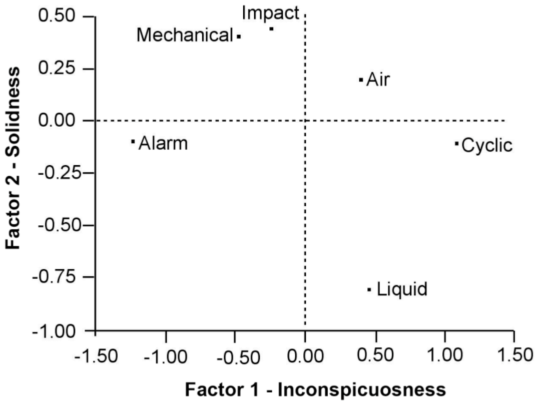


Figure 5A. Product sound categories presented as a function of Factor 1 - Inconspicuousness and Factor 2 – Solidness.

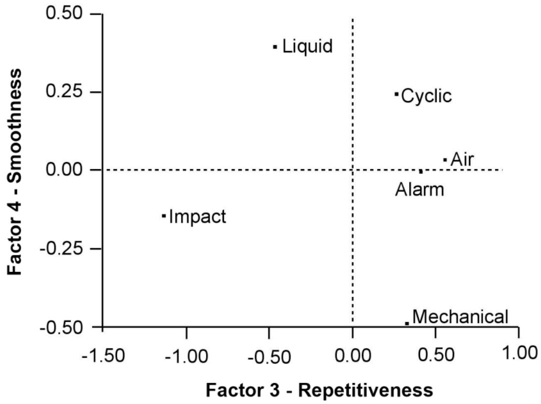


Figure 5B. Product sound categories presented as a function of Factor 3 - Repetitiveness and Factor 4 – Smoothness.

Discussion

The factor analysis of the rating attributes resulted in five factors that categorize the evaluative dimensions of product sounds. These factors are Inconspicuousness, Solidness, Repetitiveness, Smoothness, and Familiarity. The *Inconspicuousness* factor relates to the sensory judgments because the psychoacoustical adjectives loaded on this factor. Sensory judgments on a sound result in basic emotional responses (van Egmond, 2004) that may determine the attentive value of a sound. Therefore, listeners respond to some sounds more attentively than others. Alarm

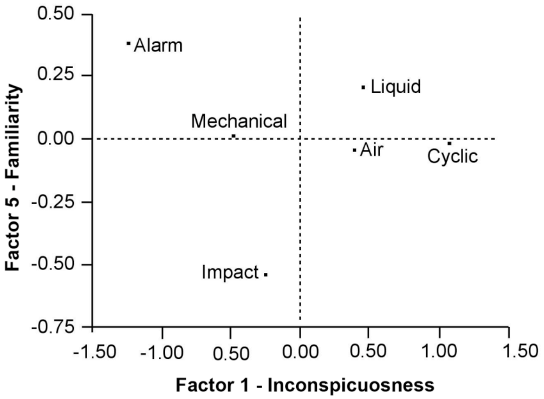


Figure 5C. Product sound categories presented as a function of Factor 1 - Inconspicuousness and Factor 5 – Familiarity.

sounds, for example, are deliberately designed to sound sharp, loud, and high pitched to evoke such sensations. Cyclic sounds, on the other hand, are a consequence of a functioning product (e.g., kitchen extractor fan) and often are regarded as background sounds.

Another factor that relates to the auditory quality of the sound is the *Smoothness* factor. Two of the constituting words of this factor (smooth, round) relate to the auditory smoothness and one of the constituting words (hollow) relates to the material quality of the sound source. In Figure 5B, it can be seen that listeners associate the auditory smoothness to material softness. Liquid, Cyclic, and Air sounds that are caused by aero- and hydro-dynamic events are perceived as smooth sounds.

The *Solidness* factor relates to the tangible (material) qualities of the sound source. This factor organizes the product sound domain into four categories based on the physical state of the interacting materials: solid, liquid, air, and digital. On this factor dimension, the sound categories seem to be clustered by the material composition of the sound source, which is similar to Gaver's classification of everyday sounds (1993a)—except Alarm sounds. Digitally produced sounds such as warning beeps of the microwave oven constitute the alarm sounds of the product sound domain. Alarm sounds are mostly digitally produced via the activation of an electric buzzer that produces auditory signals. It is hard to classify alarm sounds in terms of interacting materials. Furthermore, auditory qualities such as loudness and roughness appeared in Solidness factor. Thus, loudness and roughness perception can be related to the machinery nature of sound source.

Repetitiveness factor relates to the temporal aspects of product sounds and is a good indicator whether a sound is produced manually or electrically. Sounds produced by electrical devices (air, cyclic, machine, alarm sounds) are described as repetitive and constant. Sounds produced by users' actions are often not that regular (liquid and impact sounds).

Familiarity with the product sounds may imply if listeners are successful in sound identification. Alarm sounds appear to be the most familiar sounds and impact sounds appear to be the least familiar. The remaining categories are not distinguished by familiarity. Similarly to Ballas' findings (1993), this may indicate that sounds that have a distinct spectral-temporal structure (i.e., alarm sounds) are identified easier than sounds that are short and noisy (i.e., impact sounds).

The words that were eliminated cannot be used to describe all types of product sounds. These words are specific to certain product sound category. For example, the words 'closing' and 'wood' can describe only impact sounds. The material description 'glass' can only be used for liquid sounds. It is the alarm sounds that are digital.

To summarize, verbal attributes of products sounds relate to auditory properties, source properties, and familiarity. This finding indicates that there are sensorial judgments as well as judgments on the source of the sound.

Experiment 4

Verbal attributes of product sound categories

In Experiment 3, the semantic associations of each sound were investigated. However, do the individual members of a category evoke the same semantic associations as a category? A category may activate higher-level concepts compared to its constituent sounds, because a category represents common features of its constituents. Thus, in Experiment 4, the semantic associations of a product sound category will be investigated.

In Experiments 1 and 2 it was shown that sharpness, roughness, low-frequencies, and loudness parameters underlie the similarity judgments for the sound categories. Zwicker & Fastl (1990) proposed a model of sensory pleasantness that contains sharpness, roughness, loudness, and noisiness parameters. Bisping (1997) has suggested that car sounds can be described by two perceptual factors (power and pleasantness). The Power factor was mainly related to loudness and low-frequencies. These factors may also be relevant to the product sound domain. Similarly, In Experiment 1 relatively large-sized products were positioned high on sharpness-loudness and noisiness-low frequencies dimensions. Conversely, small-sized products (e.g., shavers, alarm clock) were positioned relatively low on noisiness –low frequencies dimension. This finding may relate to the concept of *power*. Whereas product sounds (e.g., liquid sounds) that are positioned low on sharpness-loudness dimension can be perceived as less powerful but more pleasant. Thus, this finding may relate to sensation of *pleasantness*.

Therefore, Experiment 4 uses words that describe power perception (i.e., powerful, weak, machine) and pleasantness (i.e., pleasant, obtrusive). These words replaced the some of the psychoacoustical and source property descriptions used in Experiment 3. In addition, the other attributes were identical to Experiment 3. Similar to Experiment 3, listeners were asked to rate six domestic product sound sequences on 48 semantic attributes, or to choose 'not applicable' (N/A). The procedure was identical to Experiment 3, except for the sounds presented. We created six sound sequences each of which was composed of three product sounds that represent one product sound category from Experiment 1.

Method

Participants

Thirty-three students of the Delft University of Technology participated. The average age of the participants was 21.5 years. All participants reported normal hearing. Students voluntarily participated and were paid.

Stimuli

Six sound sequences were created to represent one of the six perceptual product sound categories from Experiment 1. For each category, three of the most representative sounds were selected from Experiment 1 to create the sound sequences. The sound sequences had a maximum duration of 15 seconds (maximum 5 seconds for each sound) and were presented at a similar comfortable listening level.

Rating attributes

Forty-eight attributes (i.e., descriptive words) are presented in Table 4 as a function of basic concepts (column names). The selection of the attributes followed the same procedure as in Experiment 3. The words presented in italics in Table 4 were only used in Experiment 4. The words were translated from English language into Dutch language.

Apparatus

The stimuli and the attributes were presented by using the same software application as in Experiment 3 on a Macintosh iMac G4 700 MHz computer with a 15" screen through Sony MDR-CD550 headphones. The experiment took place in a quiet room.

Procedure

The procedure was identical to Experiment 3.

Results

Data were analyzed using the same procedure as in Experiment 3.

Elimination of attributes

In 10% of the cases participants rated ‘N/A’. Two participants with ‘N/A’ rating frequencies of 52% and 62% were excluded from further analysis. Table 7 presents the inappropriate words as a function of product sound categories and methods of analysis. The last column (All Words) indicates the total amount of words that were excluded for further analysis. In this table, the word ‘glass’ was never found appropriate to describe any sound category. The words ‘bathroom’, ‘smooth’ and ‘massive’ were found inappropriate only for alarm sounds, ‘bedroom’ and ‘manual’ for only cyclic sounds, ‘long’ and ‘electric’ for only impact sounds, ‘metal’ for only liquid sounds, and ‘relax’ for only mechanical sounds.

Factor analysis

The ratings were analyzed using the method of principal components analysis with Varimax rotation. For the analysis, ‘N/A’ ratings were replaced by the mean values taken over participants and sound sequences for each descriptive word. Five factors explained 67% of the variance. A reliability analysis using Cronbach’s alpha model was conducted in order to check the internal consistency of the attributes in a factor. The Cronbach’s alpha values of each factor are presented in Table 8. These values ranged from .85 to .42. Factor 4 had a high alpha value (.70). The words loaded on this factor also had very high communalities (familiar: .87, strange: -.85) indicating a strong consistency within the factor. The factors, the attributes, and the explained variance are presented in Table 8. The factors can be interpreted as follows:

On *Factor 1,* words ‘unpleasant, obtrusive, tense, and fast’ positively loaded high and word ‘slow’ negatively loaded high. These words are related to negative emotions and the operation speed of the product. Therefore Factor 1 was interpreted as *Unpleasantness*. On *Factor 2,* words ‘constant and repetitive’ loaded positively high and word ‘irregular’ loaded negatively high. Therefore, Factor 2 was interpreted as *Repetitiveness*. On *Factor 3,* words ‘mechanic, hard, and machine’ loaded positively high. Therefore, Factor 3 was interpreted as *Machinery*. On *Factor 4,* words ‘strange’

Analysis Type								***All Words***
Correspondence							**Variance**	
Air	**Alarm***	**Cyclic**	**Impact**	**Liquid***	**Mechanical**	***All***		
-	Air	-	Air	-	Air	*Air*	Air	*Air*
-	Bathroom	-	-	-	-	*Bathroom*	Bathroom	*Bathroom*
-	-	Bedroom	-	-	-	*Bedroom*	Bedroom	*Bedroom*
-	Blowing	-	Blowing	-	Blowing	*Blowing*	Blowing	*Blowing*
Closing	Closing	Closing	-	Closing	Closing	*Closing*	-	*Closing*
Digital	-	Digital	Digital	-	-	*Digital*	Digital	*Digital*
-	Droning	-	Droning	Droning	-	*Droning*	-	*Droning*
Edged	-	Edged	-	-	-	*Edged*	-	*Edged*
-	-	-	Electric	-	-	*Electric*	-	*Electric*
Glass	Glass	Glass	Glass	Glass	Glass	*Glass*	-	*Glass*
-	-	-	-	-	-	-	Hollow	*Hollow*
Impacting	-	Impacting	-	-	-	*Impacting*	-	*Impacting*
-	-	-	-	-	-	-	Kitchen	*Kitchen*
Liquid	Liquid	-	Liquid		Liquid	*Liquid*	-	*Liquid*
-	-	-	Long	-	-	*Long*	-	*Long*
-	-	Manual	-	-	-	*Manual*	Manual	*Manual*
-	Massive	-	-	-	-	*Massive*	-	*Massive*
-	-		-	Metal	-	*Metal*	-	*Metal*
Opening	-	Opening	-	Opening	Opening	*Opening*	Opening	*Opening*
Plastic	-	Plastic	-	-	-	*Plastic*	Plastic	*Plastic*
-	-	-	-	-	Pleasant	*Pleasant*	-	*Pleasant*
Pouring	Pouring	-	Pouring	-	Pouring	*Pouring*	Pouring	*Pouring*
-	-	-	-	-	Relaxed	*Relaxed*	-	*Relaxed*
Reserved	-	-	-	Reserved	Reserved	*Reserved*	Reserved	*Reserved*
-	Rotating	-	Rotating	Rotating	-	*Rotating*	-	*Rotating*
Rubber	Rubber	Rubber	Rubber	Rubber	Rubber	*Rubber*	-	*Rubber*
Short	-	Short	-	-	-	*Short*	-	*Short*
Small	-	Small	-	-	-	*Small*	-	*Small*
-	Smooth	-	-	-	-	*Smooth*	Smooth	*Smooth*
-	-	-	Soft	-	Soft	*Soft*	-	*Soft*
Wood	Wood	Wood	-	-	Wood	*Wood*	-	*Wood*
Total number of words								
14	13	13	12	8	13	*29*	13	*31*

* The words in these categories were clustered only with 'N/A' ratings (not together with '1' ratings)

Table 7. Eliminated attributes in Experiment 4 for each product sound category (stream) in each analysis.

loaded positively high and word ‘familiar’ loaded negatively high. Therefore, Factor 4 was interpreted as *Unfamiliarity*. On *Factor 5* Words ‘powerful and big’ loaded positively high and word ‘weak’ loaded negatively high. These words indicate the power that is employed to operate a product. Therefore, Factor 5 was interpreted as the *Power.*

The mean of the regression weights of the factor scores were determined for each sound sequence. In Figures 6A, B, and C, the sound sequences are presented in three factor spaces. (In order to prevent abundance of data presentation, we have chosen to present only three combinations of factor dimensions). According to the figures, *air* sounds are positioned the highest on the Repetitiveness factor and rather high on the Power factor. *Alarm* sounds are positioned the highest on the Unpleasantness factor and the lowest on Unfamiliarity factor. *Cyclic* sounds are positioned the highest on Power factor and the lowest on the Unpleasantness factor. *Impact* sounds are positioned the lowest on the Repetitiveness and rather high on the Unfamiliarity factor. Liquid sounds are positioned the lowest on both the Power factorand Machinery factor. *Mechanical* sounds are positioned the highest on both the Mechanical and the Unfamiliarity factors, and rather high on the Unpleasantness factor.

Comparison of the factors of Experiment 3 and Experiment 4

Although the same sounds were used for Experiment 3 and 4, the sounds were presented separately in Experiment 3 and as a sequence in Experiment 4. Common attributes were used in both experiments—except the 10 attributes that were used for psychoacoustics in Experiment 3 and another 10 for power perception and pleasantness in Experiment 4. In order to allow a comparison between the factors of the two experiments, the mean of the regression weights for each product sound category of Experiment 3 was calculated. These mean regression weights were correlated with those of Experiment 4.

Table 9 presents the correlations of the each factor weights from Experiment 3 and from Experiment 4. Inconspicuousness factor of Experiment 3 and Unpleasantness factor of Experiment 4 exhibited a strong and negative correlation ($r(4)=-.98$, $p<.0001$). Smoothness factor of Experiment 3 was negatively correlated with the Solidness factor of Experiment 3 ($r(4)=-.84$, $p<.05$) and Machinery factor of Experiment 4 ($r(4)=-.89$, $p<.05$). Solidness factor of Experiment 3 was positively correlated with Machinery factor of Experiment 4 ($r(4)=.89$, $p<.05$). Another high and negative correlation was observed for the Familiarity factor of Experiment 3 and Unfamiliarity factor of Experiment 4 ($r(4)=-.83$, $p<.05$). Repetitiveness factors of Experiment 3 and Experiment 4 were highly correlated ($r(4)= .92$, $p<.05$).

	Factor loading				
Attributes	**1**	**2**	**3**	**4**	**5**
Obtrusive	**.77**	.15	.35	.01	.18
Tense	**.77**	.10	.36	.09	.09
Fast	**.74**	.02	.00	-.02	.06
Unpleasant	**.72**	.08	.36	.06	.05
Slow	**-.70**	-.05	.16	.05	.02
Constant	-.02	**.81**	-.06	.01	.06
Irregular	-.17	**-.71**	.04	.21	.12
Repetitive	.26	**.63**	.33	-.26	.11
Mechanic	.08	.03	**.83**	.12	.05
Hard	.39	-.32	**.66**	.00	.03
Machine	.00	.46	**.62**	.04	.24
Familiar	.08	.11	.01	**-.85**	.00
Strange	.13	-.13	.16	**.84**	-.04
Powerful	.21	.04	.18	.01	**.81**
Big	-.43	.31	.05	.09	**.60**
Weak	-.37	.32	.04	.23	**-.60**
% of variance	*21.17*	*13.48*	*13.04*	*10.19*	*9.43*
α	*.85*	*.70*	*.64*	*.70*	*.42*

Note. Boldface indicates highest factor loadings > .04. Factor 1 = Unpleasantness, Factor 2 = Repetitiveness, Factor 3 = Machinery, Factor 4 = Unfamiliarity, Factor 5 = Power.

Table 8. Attributes, Factor Loadings a Five-Factor Solution, Percentages of Variance Explained, and Cronbach's Alpha for the Attributes of Experiment 4.

Discussion

Experiment 4 investigated whether semantic associations of sounds would differ if the sounds were presented as a category (sound stream) instead of as individual sounds as in Experiment 3. The factor analysis of the rating attributes resulted in five factors. These factors are Unpleasantness, Repetitiveness, Machinery, Unfamiliarity, and Power. Of these five factors, two of them are similar to the factors of Experiment 3: The Repetitiveness and Unfamiliarity factors are similar to the Repetitiveness and the Familiarity factors of Experiment 3. The *Unfamiliarity* factor contains the same attributes as the Familiarity factor in Experiment 3 (familiar, unfamiliar). The *Repetitiveness* factor contains two attributes (repetitive and constant) that are the same to attributes that constitute Repetitiveness factor in Experiment 3. Furthermore, the *Machinery* factor shares two common attributes (hard and mechanic) with the Solidness factor of Experiment 3. However, the Machinery factor can be distinguished from the Solidness factor because this factor relates more to machinery sounds as mechanical sounds are positioned high on the factor dimension.

The occurrence of *Unpleasantness* and *Power* factors somewhat relate to Bisping's (1997) findings and was as predicted. The *Unpleasantness* factor can be explained by the acoustical content of the sound. As proposed by Zwicker and Fastl (1990) unpleasant experiences with respect to auditory stimuli may occur as a result of a sensory judgment that is determined by the perceived auditory quality of the sounds (e.g., sharp, rough, loud, noisy). The current experiments have also demonstrated that product sounds are machinery sounds that can be perceived as loud, rough, and noisy. Thus, it is expected that listeners judge product sounds on an emotional level. In addition, this factor contains attributes (fast, slow) that are related to the temporal aspect of the sounds. As a result, Alarm and Mechanical sounds are the mostunpleasant sounds. This may be partly caused by their spectral-temporal content and partly by the higher attentive value of the high-pitched sounds. Thus, unpleasantness on a cognitive level can also occur, for example, if a listener cannot intervene a mechanical sound (e.g., shaver sound).

Bisping (1997) has discussed that the perceived power can be a result of the amount low-frequencies in the spectrum of a sound and can be a good indicator for the (car) engine performance. This indicates that although power judgment depends on the spectral content of a sound, it is a fairly cognitive judgment. Similarly, *Power* factor contains attributes that related to the sound source and its size. Consequently, such a judgment relates to operating capacity of the product. For example, big products (e.g., washing machine) require more power to operate and produce lower frequency sounds and small products (e.g., toothbrush), the opposite.

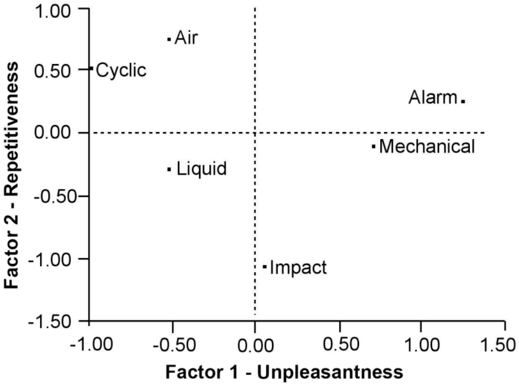


Figure 6A. Product sound categories presented as a function of Factor 1-Unpleasantness and Factor 2 – Repetitiveness.

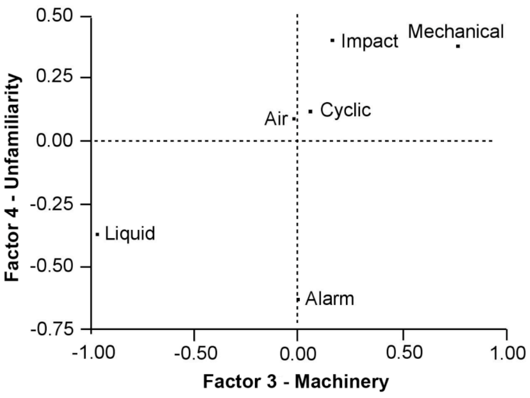


Figure 6B. Product sound categories presented as a function of Factor 3 - Machinery and Factor 4 – Unfamiliarity.

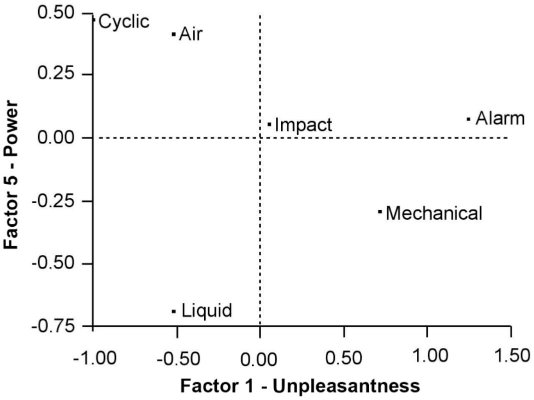


Figure 6C. Product sound categories presented as a function of Factor 1 - Unpleasantness and Factor 5 – Power.

Compared to Experiment 3, this experiment has resulted in higher-level concepts. For example, Solidness factor (and somewhat Smoothness factor) of Experiment 3 concerns the material properties of the sound source, whereas Machinery and Power factors of Experiment 4 concern the type of sound source and its capacity to operate. Furthermore, Inconspicuousness factor of Experiment 3 concerns psychoacoustical attributes (i.e., perceived auditory quality). However, Unpleasantness factor of

Experiment 4 indicates listeners' (emotional) response to these attributes. Other factors occurred to be very similar in both experiments (i.e., Repetitiveness and Unfamiliarity factors). With the comparison of the factors of Experiments 3 and 4, we can conclude that semantic judgments concerning individual sounds may resemble the semantic judgments of a sound category. However, sound categories may be related to higher-level concepts and can be on a cognitive and emotional level.

	Experiment 3				
Experiment 4	**Inconspicuousness**	**Solidness**	**Repetitiveness**	**Smoothness**	**Familiarity**
Unpleasantness	-.98	.32	.14	-.64	.29
Repetitiveness	.26	-.13	.92	.24	.56
Machinery	-.35	.89	.29	-.89	-.32
Unfamiliarity	.30	.72	-.25	-.55	-.83
Power	.22	.48	.38	-.01	-.26

Note. Correlations are taken over six sound categories (*N* = *6*)

Table 9. Correlation matrix of regression weights from factor analysis of Experiments 3 and 4.

Experiment 5

Constituents of similarity judgments

The criteria for category formation and the underlying processes of categorization depend heavily on the strength of similarity between objects (Medin & Barsalou, 1987; Medin, Lynch, & Solomon, 2000; Rosch, 1978). Especially, Eme & Marquer (1998) have distinguished five main groups of strategies that occur during a visual-similarity comparison task. These strategies are holistic, analytical, partial, one-feature, and shift strategies. In the *holistic* strategy, the shapes are encoded and compared as a whole. In the *analytic* strategy, a set of basic features and mental image of individual units are stored and compared. In the *partial* strategy, the shape is encoded as a whole and some features are stored. In the *one-feature* strategy, only one 'critical' feature or unit is encoded and compared. The *shift* strategy indicates that a subject alters between two or three strategies on a single comparison.

For product sounds, similar criteria may apply. First, two sounds may *acoustically* sound similar due to the similarity in the spectral-temporal structure (e.g., mechanical alarm clock and kitchen timer sounds). Secondly, if structural comparisons are absent, then similarity judgments can be determined on a semantic or a conceptual level. Two sounds may be judged as similar because they *semantically* refer to the same event or object (e.g., alarm clock sound, digital or mechanical) or because they are *conceptually* related (the sound of the rinse cycle of a washing machine and the sound of the centrifuge cycle of a washing machine both refer to a clothes washing activity). Furthermore, the study of Aldrich et al. (in press) supports this assumption that environmental sounds are found similar not only because of their acoustical features, but also because of the objects that cause the sounds.

Experiment 5 investigates further the underlying processes in similarity judgments. The similarity judgments in Experiment 5 were obtained using a paired-comparison paradigm. A list of concepts that could be the components of the strategies for similarity judgments was given to the participants so as to indicate the most frequently used strategies. This list mainly contained the basic concepts by which product sounds are represented (see Experiment 1 and 2): action, emotion, location, material, abstract meaning, onomatopoeia, psychoacoustics, sound type, source, source properties, and temporal structure.

With Experiment 5, we also wanted to see whether the product sound categories in Experiment 1 and Experiment 2 did not occur as an artifact of the used paradigm (i.e., free categorization). Perhaps, if category borders remain the same after using another paradigm (paired-comparisons), then the plausibility of the categories will be confirmed.

Method

Participants

Eighty students of the Delft University of Technology participated. The average age of the participants was 22 years. All participants reported normal hearing. Students voluntarily participated and were paid.

Stimuli

Nineteen product sounds that shared high acoustical similarities were selected to represent four sound categories defined in Experiment 1. These categories were air, cyclic, liquid, and mechanical sounds. The sound recordings of various electrical domestic appliances in operation were used as stimuli. The sounds were edited on a Macintosh PowerPC G4 computer using the sound-editing program Sound Studio (version for Mac OS X), sounds longer than 5 seconds were trimmed to a maximum duration of 5 seconds. All sounds were saved in a stereo format with a sampling rate of 44.1 kHz and 16 bits. The loudness levels were adjusted to a comfortable listening level for each sound. The participants were not allowed to change the sound levels during the experiment.

Apparatus

The stimuli and the descriptive words were presented by a specially designed software application on a Macintosh PowerBook G4 computer with a 12" screen through Sony MDR-CD550 headphones. The experiment took place in a quiet room.

Procedure

Before the study started, each participant received a brief explanation about the purpose of the study on an A4 sized paper. The experiment had two phases. In the first phase, perceptual similarities were rated based on a paired comparison task; in the second phase, the most frequently used strategies (product sound related basic concepts) were rated on a questionnaire. Participants were not told about the second phase of the experiment.

In the similarity judgment task, a participant received 48 sound pairs out of 190 pairs—they took on average 15 minutes to judge. Two sounds were randomly selected for each pairwise sound presentation. Each pair was presented 20 times using a Monte Carlo method (Press, Teukolsky, Vetterling, and Flannery, 1995). A participant received the sound pairs on a screen as two different sound buttons on which it was written ‘sound 1’ and ‘sound 2’. The order of the sound pairs and the order between the sounds in a pair were randomized for each participant. Then, a participant listened to each of the sounds and rated their similarity on a 6-point scale (1 - not similar, 6 – very similar). This was repeated for all 48 sound pairs.

A questionnaire followed the similarity judgment task. A participant answere the question “on what bases have you found similarities between the sound pairs?” and the following 11 product sound related basic concepts were provided as options for their answers: action, emotion, location, material, meaning, onomatopoeia, psychoacoustics, sound type, source, source properties, and temporal. Next to each basic concept a couple of examples were provided to facilitate the participants’ decision (e.g., for material ‘plastic, metal, etc.’ or for temporal aspects ‘continuous, repetitive, multiple, single, constant, etc.). A participant indicated the frequency by which they sought similarity between the sound pairs on a 5-point scale (1 indicating never, 5 indicating always) each of the descriptive groups.

Results

Similarity judgment

It was determined if the sounds in the pair belonged to the same sound category or to different sound categories. The sounds that belonged to the same sound category were labeled as ‘similar’ (e.g., a shaver and a toothbrush sound both belong to the mechanical sound category, therefore they are similar). The sounds that belonged to two different sound categories were labeled as ‘dissimilar’ (e.g., a shaver and a hairdryer sound respectively belong to the mechanical sound category and the air sound category, therefore they are dissimilar). Thus, the similarity ratings were averaged over the ‘similar’ and ‘dissimilar’ sound pairs. Same category sound pairs (*M* = 3.55, *SE* = .07) were rated significantly higher than dissimilar sound pairs (*M* = 2.10, *SE* = .09), F (1, 78) = 481.91, *p* < .001.

An additional analysis was conducted in which the dissimilar sound pairs were differentiated by the sound categories they belonged to. The following seven pairs resulted: air-cyclic, air-liquid, air-mechanical, cyclic-liquid, cyclic-mechanical, liquid-mechanical. For example, if the pair consisted of a shaver (mechanical sound) and a hair dryer sound (air sound), the label ‘mechanical-air’ is assigned to this dissimilar sound pair. This was done for all possible combinations. The sound pairs that contained the same sound categories but differed in the order of sound presentation (e.g., air-mechanical vs. mechanical-air) were treated equally. The sound pairs that contained sounds from one sound category (e.g., mechanical-mechanical) were labeled as ‘similar’. The similarity ratings were averaged for each dissimilar sound pair (air-cyclic, air-liquid, air-mechanical, cyclic-liquid, cyclic-mechanical, liquid-mechanical) and the similar sound pairs (e.g., air-air, liquid-liquid, etc.). Figure 7 presents the mean similarity ratings as a function of the dissimilar sound pairs and similar sound pairs. According to the figure, the similar sound pairs had the highest similarity rating (3.55). Among the dissimilar sound pairs, the air-cyclic sound pair had the highest similarity rating (3.06) followed by air-mechanical (2.01), cyclic-mechanical (2.00) and cyclic-liquid (1.95). The air-liquid sound pair had the lowest similarity rating (1.66) followed by the liquid-mechanical sound pair (1.86).

The averaged similarity ratings per participant were analyzed by an ANOVA with similarity type as the within-subjects factors (7 levels). A significant effect was found for the similarity type, *F*(6, 468) = 147.58, *p*<.001. A planned comparison was conducted to determine which sound pairs differed significantly. Similar sound pairs differed significantly from the dissimilar sound pairs (*p*<.001). Air-cyclic sound pairs were differed significantly from other dissimilar sound pairs (*p*<.001). In addition, air-liquid differed significantly from cyclic-liquid (*p*<.05) and differed significantly from air-cyclic, air-mechanical, cyclic-liquid, and cyclic-mechanical (*p*<.001).

Strategy ratings

The strategy ratings were averaged over the sound descriptive groups. The mean values for the strategy ratings are: psychoacoustics (*M* = 3.87, *SE* = .13), action descriptions (*M* = 3.65, *SE* = .13), onomatopoeias (*M* = 3.65, *SE* = .14), sound source descriptions (*M* = 3.11, *SE* = .16), sound type (*M* = 3.06, *SE* = .15), temporal descriptions (*M* = 3.03, *SE* = .15), emotion (*M* = 2.62, *SE* = .15), material (*M* = 2.16, *SE* = .14), abstract meaning (*M* = 2.10, *SE* = .14), location descriptions (*M* = 2.06, *SE* = .13), and source property descriptions (*M* = 2.01, *SE* = .14).

A principal component analysis (PCA) with Varimax rotation was conducted on the strategies (source, action, psychoacoustics, etc.). The PCA resulted in three factors which explained 53% of the data. The resulting factors and the explained variance for each factor are presented in Table 10. The factors are interpreted as follows.

On *Factor 1* factor, mainly the sound descriptive groups that are related to the semantic representation of sounds loaded (source, location, source properties, action, emotion, and meaning). Therefore, this factor was interpreted as *Cognitive Evaluation*. On *Factor 2*, mainly the sound descriptive groups that are related to the spectral-temporal structure of a sound loaded (temporal, onomatopoeia, psychoacoustics). Therefore, this factor was interpreted as *Perceptual Evaluation*. On *Factor 3*, material and sound type loaded, which both relate to the sound quality and sound source. Therefore, this factor was interpreted as *Associative Evaluation*.

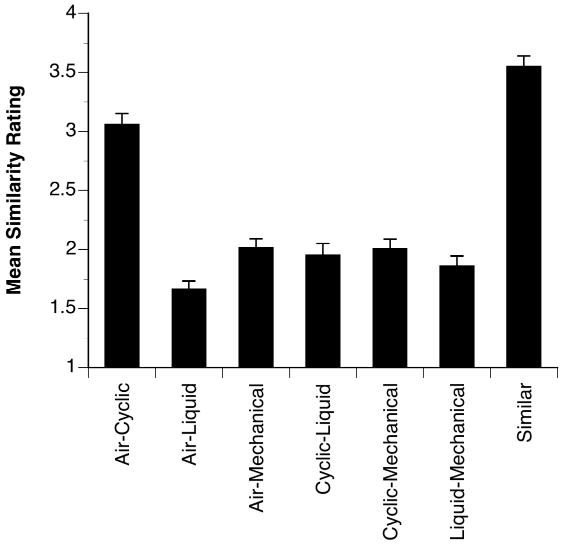


Figure 7. *Mean similarity ratings of sound pairs of dissimilar and similar sound categories.*

Discussion

Experiment 5 has resulted in two main findings. First, product sounds are perceptually similar within a category and dissimilar between categories. Only air and cyclic sounds are perceived as similar sound categories. Second, we found three types of evaluation on which similarity judgments were based (i.e., cognitive, perceptual, and associative evaluation). Especially the perceptual evaluation received the highest ratings and, therefore, seems to be the most frequently used evaluation for similarity judgments. Although not so frequent, similarity judgments may also be based on cognitive evaluations.

Similarity within and dissimilarity between product sound groups

Product sound categories seem to be perceptually distinguishable. Although equal sounds were never compared, the average rating for the pairs of sounds within a category was still higher than the mid-point of the rating scale. This indicates that sounds are sufficiently similar to be in the same category. In addition, the results also provide converging evidence that the category borders found in Experiment 1 are perceptually salient.

	Factor loading		
Strategies	**1**	**2**	**3**
Source	**.718**	-.065	.066
Location	**.631**	-.034	.232
Source Properties	**.626**	.056	.139
Action	**.612**	.283	.171
Emotion	**.562**	.047	-.403
Meaning	**.490**	.015	.382
Temporal	.063	**.783**	.041
Onomatopoeia	.162	**.748**	-.043
Psycho-acoustics	-.114	**.734**	.059
Material	.111	.089	**.814**
Sound Type	.280	-.014	**.730**
% of variance	*21.54*	*16.47*	*14.70*

Note. Boldface indicates highest factor loadings > .04. BW = band width. Factor 1 = Cognitive Evaluations, Factor 2 = Perceptual Evaluations, Factor 3 = Associative Evaluations.

Table 10. Strategies, Factor Loadings for a Three-Factor Solution, and Percentages of Variance Explained for Similarity Judgments of Experiment 5.

Good dissimilarity across category members is essential for clear borders between categories; and it also ensures that categorical interpretations are not mistaken for one sound (Mervis & Rosch, 1981). However, product sounds having high structural similarities within a category will be treated not only perceptually but also cognitively similar. Although such sounds may still evoke similar conceptual associations (e.g., air sound, loud, house); it may be hard to lexically distinguish the sounds within a category (e.g., a hairdryer sound and a vacuum cleaner sound). On that account, it may even be harder to assign a correct category for structurally similar sounds from different categories (e.g., Air and Cyclic sounds). Because, this time activated concepts and lexical representations will belong to two different categories, and thus, will be too many. These two situations both exemplify how ambiguity may occur as a result of perceptual similarity both within and across category members.

Previous literature has discussed that free categorization paradigm results in semantically similar sound categories and paired-comparison paradigm results in acoustically similar sound categories (Aldrich et al., in press; Gygi et al., 2007). That is, not only similarity judgments were based on different strategies but also occurring categories were different. Our results showed that although the employed strategies for similarity judgments were different with respect to the paradigm used, categorical relationship remained the same.

One explanation for this may be that in our study we focused on one sub-category of environmental sounds that does not have a wide range of categorical associations (unlike the sounds used in the studies of Aldrich et al. (in press) and Gygi et al. (2007)). Thus, acoustical similarity between product sounds may have yielded only a limited number of common concepts. Another explanation would be that for the product sounds we employed acoustical similarity within a group was more salient because sound sources and actions causing the sounds were similar within a category. Thus, active comparison of product sounds may have also been affected by the automatic activation of the conceptual associations (see Orgs, Lange, & Dombrowski, 2006). In conclusion, product sounds that are acoustically similar are conceptually similar too. Thus, the structural composition of a product sound can be indicative of its categorical differentiation.

Strategies for similarity judgments

The results indicate that there are three main factors that underlie the similarity judgments for the sound categories. These can be explained as cognitive, perceptual, and associative evaluations. *Cognitive evaluations* are related to the cause of the sound (source, action, and the source properties), the location in which the sound can be heard, and the abstract meaning. When similarity judgments are based on a cognitive evaluation, listeners' use knowledge that results from a previous experience with a product. The emotional judgments are also a part of this factor. This is inline with general theories of emotion that claim that certain emotional judgments (not basic affects which derive from sensory judgments) are a result of cognitive processing (e.g., Ortony, Clore, & Collins, 1988). *Perceptual evaluations* relate to the auditory properties of a sound (psychoacoustical, temporal, and onomatopoeia). These strategies imply that the judgments have probably been based on the structural features of sounds (e.g., how repetitive or sharp the sounds are). *Associative evaluations* also underlie the similarity judgment. This finding implies that if no-recognition occurs with regard to the sound source, then, listeners base their similarity judgment on specific features of the sound or sound source (e.g., electric sounds, water sounds).

Summary of the findings

In this study the domain of product sounds and its constituent categories have been determined. Listeners distinguish six product sound categories, which are *air, alarm, cyclic, impact, liquid,* and *mechanical* sounds. Each sound category has been classified in terms of acoustical properties and their perceptual correlates. Sharpness, loudness, and noisiness are the main psycho-acoustical parameters on which the product sounds vary. Temporal constancy and repetitiveness underlie the temporal structure of the product sound categories. Within a product sound category (i.e., air, cyclic, and mechanical sounds), sounds are distinguished by the perceived roughness, calmness, loudness, and hardness of the sounds.

Semantic associations of product sounds can be structured in nine basic concepts. Labels given for sound categories indicate nine basic concepts that represent product sounds: *sources* and *actions* that cause sound, *emotions* evoked by sounds, *locations* in which sounds can be heard, *abstract meanings* that sounds convey, imitations of sounds as *onomatopoeias*, perceptual correlates of sounds such as *psychoacoustics* and *temporal descriptions*, and *sound type* caused by the materials involved in sound production. Some of these basic concepts relate specifically to certain product sound category (e.g., *meaning* for alarm sounds, *source descriptions* for liquid sounds).

Specific verbal attributes that are common to all product sounds from six categories have been determined. These attributes have been classified for individual sounds as the following factors: *inconspicuousness, solidness, repetitiveness, smoothness*, and *familiarity*, and for product sound categories as the following factors: *unpleasantness, repetitiveness, machinery, unfamiliarity*, and *power*. Comparing the factors of individual sounds and the categories showed that some of these factors are associated. For example, the pleasantness of a product sound category is associated with the inconspicuousness of the individual sounds. These factors resemble the basic concepts for product sounds. In addition, certain product sound categories are judged more familiar to listeners (e.g., alarm sounds) than others (e.g., impact sounds).

The categories stemming from a free categorization task have been confirmed by using a pair-wise comparison task in which the similarity of the sounds within and dissimilarity between categories was tested. The similarity within a category was higher than the similarity between categories. This suggests that found product sound categories are not an artifact of the employed paradigm. In addition, basic strategies on which listeners base their similarity judgment have been determined. These strategies reflect listeners' perceptual, cognitive, and associative evaluations of the sounds.

Product sound categories

In summary, the characterizing acoustical properties and semantic attributes of each product sound category are the following. *Air* sounds are relatively loud and noisy. 'Powerful' is a good descriptor for this category. They primarily consist of low frequencies in their spectral content and are constant in their temporal structure. Listeners associate these sounds to their cause (source - action) and the location in which they occur. They are similar to cyclic sounds and dissimilar to liquid sounds. *Alarm* sounds are the sharpest and the least noisy sounds, and they are repetitive. Attribution of meaning to such abstract sounds is essential because they convey meaning in certain contexts. They are conspicuous and unpleasant. *Cyclic* sounds are relatively high in loudness and less noisy compared to air sounds. In addition, they contain low frequencies in their spectral content. They are inconspicuous and are not unpleasant. Despite being inconspicuous, they are associated with locations like bathroom and kitchen. These sounds are judged the most powerful. *Impact* sounds are noisy sounds with a short duration and are unfamiliar to listeners. Listeners derive the material composition and the action from the sound. *Liquid* sounds consist of relatively low frequencies and have the lowest loudness values. They evoke pleasantness and are perceived smooth and not powerful. Listeners are familiar with such sounds and associate them to their cause and the location in which they occur. *Mechanical* sounds are relatively sharp, loud, and not so noisy. These sounds are associated with solidness and machinery. They are conspicuous and evoke a sense of unpleasantness.

General discussion

The conceptual network for product sounds seem to be organized around sound source information and auditory features of a sound. Two types of mentalrepresentations have been determined that can be functionally different. These are sensorial judgments and meaningful conceptual associations. Sensorial judgments are a result of a perceptual process, whereas the conceptual attribution occurs more on a cognitive level. Describing a product sound as loud, rough, short, or continuous is related to the auditory features of the sound. A perceptual attribute such as loudness could also be a concept. For example, a vacuum cleaner can be labeled as a 'loud' object. Such a concept can be considered on a super-ordinate level that relates to various objects (vacuum cleaner, alarm clock, and shaver).

The afore-mentioned conceptual and sensorial judgments of a sound resemble the similarity judgments that result from holistic and decompositional (analytic) strategies (Eme & Marquer, 1998) or generic knowledge and sensory perception categories (Medin & Barsalou, 1987). We have shown that processing of product sounds can remain at the sensory level or can lead to a perception of object with conceptual associations. Thus, both top-down and bottom-up process may take place for attributing meaning to a product sound. This study cannot explain when and how these processes take place. However, an interpretation can be made similar to Vanderveer's (1979). That is, if a sound is identified, source information becomes important and if no identification occurs, then the spectral-temporal structural of the sounds is available for describing the auditory percept.

Implications for product sound identification

This study is confined to the categorization of product sounds and the conceptual and semantic correlates of the categories. The results may be used to understand the underlying perceptual and cognitive functions of product sound identification. One of the main findings is that a product sound often activates concepts that relate directly to the sound source information (source, source properties, action, location, material composition, etc.). This finding is inline with Yost's (1991) propositions that the acoustic composition of a sound is converted into an auditory image that has a direct link to the concept and imagery of a sound source. Our finding together with Yost's propositions may indicate that the process of sound identification not only operates in the auditory system but also in the visual system (see also Kubovy & Valkenburg, 2000). For example, source and action descriptions may indeed relate to a visual event (e.g., coffee pouring or door closing), and location descriptions may relate to a visual scene that has a typical relation to a product or its sound (e.g., a microwave oven in a kitchen scene). Memory related studies also support the inter-connectedness of visual and auditory memory, however, on a conceptual level (e.g., Thompson & Paivio, 1994). The strong association to the sound source may originally stem from the dual coding of the concept of the product together with its auditory and visual properties (see Paivio, 1991). Thus, whenever the concept is active, the strongly associated items will be active too.

The occurrences of concepts that are not related to the sound source (e.g., psychoacoustics, temporal, emotion, location) indicate that listeners do not always access the lexical representation of a sound. These concepts may be the result of an incomplete identification process. Figure 8 presents the types of sound descriptions that may occur on stages during a sound identification process. In principle, the identification process starts with the perceptual analysis of a sound. When the perceptual analysis is completed, information about the featural aspects of a sound will be available. At this stage, psycho-acoustical and temporal descriptions (e.g., sharp, loud, repetitive, long, unpleasant) may occur due to the availability of the structural features. These auditory features are later used in the recognition phase. In other words, recognition is matching the auditory features to previously stored auditory codes. Thus, recognition always precedes identification. Identification relies heavily on the attribution of meaning. Semantic and conceptual information can be derived from a sound even before a lexicon is activated (Cummings et al., 2006). At this stage conceptual identification may occur, e.g., location or action descriptions. Theories (Fabiani et al., 1996; Vanderveer, 1979) suggest that environmental sound identification results in source description (lexical representation). We, however, suggest that the result of product sound identification depends on the sound type.

Each product sound category seems to be acoustically and semantically distinguished from another. Some sounds are characterized better by their source information (e.g., liquid sounds) and action causing the sound (e.g., impact sounds), whereas others are by abstract meanings (e.g., alarm sounds) or location (e.g., air sounds). Thus, one cannot expect that similar mechanisms operate for all types of sounds during identification. For example, identifying an alarm sound would require a direct access to semantic associations, however, for air and cyclic sounds the concept of the source needs to be activated. For impact and liquid sounds, identification of an event may be required. This suggests that the identification process may be completed at different levels for different types of product sounds.

Assuming that the identification process is completed with a sound source label (e.g., a vacuum cleaner), it does not necessarily imply that the label is correct. High perceptual similarity among sounds within and between certain categories (especially air and cyclic sounds) can create confusions in labeling. This could mean that one sound activates several sound sources, consequently, incorrect identification may result.

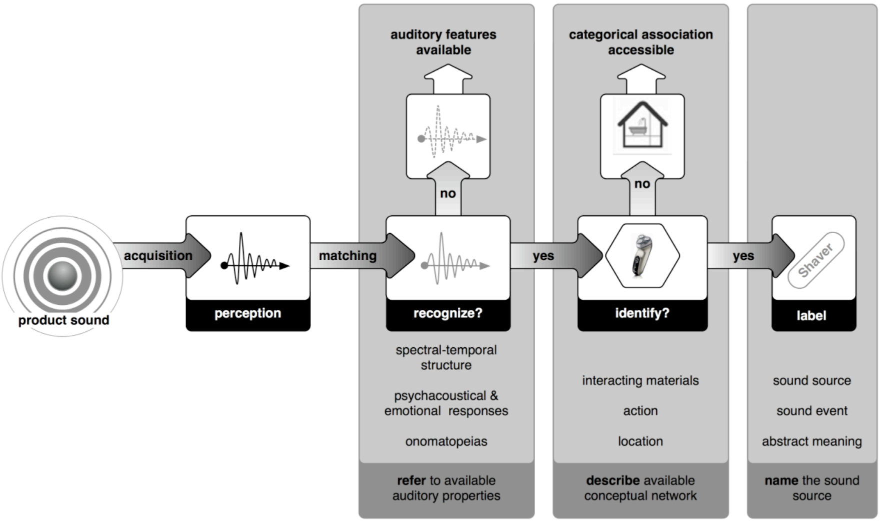


Figure 8. Labeling as a result of product sound identification.

Conclusion and future studies

This study has demonstrated that there are distinct perceptual categories for product sounds. We have suggested that the identification process for product sounds consists of both perceptual and cognitive evaluation of sounds. In addition, we have argued that the type of semantic associations is dependent on the stage of the identification process. Perceptual features of a sound are available to a listener even if the identification process has been completed. Consequently, if these features are very salient to a listener, the listener may produce semantic / conceptual associations that reflect these features and will not label the sound with a product name.

We have presented an overview on the conceptual associations that product sounds may have. In this paper, the lexical representation of product sounds has emerged as one of the concepts listeners identify a sound with, but has not been thoroughly investigated. A more specific study on how well product sounds are lexically identified is still needed. We have also argued that the produced descriptions of product sounds are a result of different stages in an identification process. Next studies may investigate more systematically the relation between semantic associations and the stages of the identification process.

Considering that both auditory and source-related visual information constitute the conceptual network of product sounds; it seems plausible that audio-visual interactions occur during sound identification. However, to what extent visual information has an additive, complimentary and/or inhibitory influence of product sound identification needs to be investigated further.

Author Note

We gratefully acknowledge the help of Esther Jansen, Anna Ruiter, Joline Boschman, Linda Rindertsma, Floris Schrader, Tim Pellikaan, and Rob Langendijk for running of the experiments. We also feel obliged to thank Alexandre Vieira for voluntarily programming the experimental software on his free time.

Footnotes

[1] Praat is a free software program for acoustical analysis for phonetics. Paul Boersma and David Weenink have implemented it, www.praat.org.

[2] Psysound is a psycho-acoustical analysis program (http://farben.latrobe.edu.au/mikropol/volume5/cabrerad/PsySound.html). For reliable measuring, it was calibrated by the SPLs of each sound for the analysis of the psycho-acoustical parameters.

References

Aldrich, K. (2005). *Perceived similarity between complex sounds: The contribution of acoustic, affective and categorical features.* Unpublished doctoral dissertation, University of Plymouth, Plymouth, UK.

Aldrich, K., Hellier, E., and Edworthy, J. (in press). What determines auditory similarity? The effect of stimulus group and methodology. *Quarterly Journal of Experimental Psychology.*

Aures, W. (1985). Berechnungsverfahren für den sensorischen Wohlklang beliebiger Schallsignale. *Acustica*, *59*, 130-141.

Ballas, J. A. (1993). Common factors in the identification of an assortment of brief everyday sounds. *Journal of Experimental Psychology: Human Perception and Performance, 19*(2), 250-267.

Ballas, J. A., & Mullins, T. (1991). Effects of context on the identification of environmental sounds. *Human Performance, 4*(3), 1991-1219.

Barsalou, L. W. (1999). Perceptual symbol systems. *Behavioral and Brain Sciences, 22*(4), 577-609.

Bartlett, J. C. (1977). Remembering environmental sounds: Role of verbalization at input. *Memory & Cognition, 5*(4), 404-414.

Benko, U., Petrovcic, J., Juricic, D., Tavcar, J., Rejec, J., & Stefanovska, A. (2004). Fault diagnosis of a vacuum cleaner motor by means of sound analysis. *Journal of Sound and Vibration, 276*(3-5), 781-806.

Biederman, I. (1987). Recognition-by-components: a theory of human image understanding. *Psychological Review, 94*(2), 115-147.

Bisping, R. (1997). Car interior sound quality: Experimental analysis by synthesis. *Acustica, 83*(5), 813-818.

Björk, E. A. (1985). The perceived quality of natural sounds. *Acustica, 57*(3), 185-188.

Blauert, J., & Jekosch, U. (1997). Sound-quality evaluation: A multi-layered problem. *Acustica, 83*(5), 747-753.

Borghi, A. M., Caramelli, N., & Setti, A. (2005). Conceptual information on objects' locations. *Brain and Language, 93*(2), 140-151.

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound.* Massachusetts: MIT Press.

Cleary, A. M. (2002). Recognition with and without identification: Dissociative effects of meaningful encoding. *Memory & Cognition, 30*(5), 758-767.

Cummings, A., Ceponiene, R., Koyama, A., Saygin, A. P., Townsend, J., & Dick, F. (2006). Auditory semantic networks for words and natural sounds. *Brain Research, 1115*(1), 92-107.

Deutsch, D. (1980). The processing of structured and unstructured tonal sequences. *Perception & Psychophysics, 28*(5), 381-389.

Deutsch, D., & Feroe, J. (1981). The internal representation of pitch sequences in tonal music. *Psychological Review, 88*(6), 503-522.

Edworthy, J., Hellier, E., & Hards, R. (1995). The semantic associations of acoustic parameter commonly used in the design of auditory information and warning signals. *Ergonomics, 38*(11), 2341 - 2361.

Eme, P. E., & Marquer, J. (1998). Quantitative and qualitative individual differences in visual comparison processing: Are strategies only analytic or holistic? *Personality and Individual Differences, 25*(3), 459-475

Fabiani, M., Kazmerski, V. A., Cycowicz, Y. M., & Friedman, D. (1996). Naming norms for brief environmental sounds: Effects of age and dementia. *Psychophysiology, 33*(4), 462-475.

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

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

Gygi, B., Kidd, G. R., & Watson, C. S. (2004). Spectral-temporal factors in the identification of environmental sounds. *Journal of the Acoustical Society of America, 115*(3), 1252-1265.

Gygi, B., Kidd, G. R., & Watson, C. S. (2007). Similarity and categorization of environmental sounds. *Perception & Psychophysics, 69*(6), 839-855.

Jekosch, U. (1999). Meaning in the context of sound quality assessment. *Acustica, 85*(5), 681-684.

Kendall, R. A., & Carterette, E. C. (1993). Verbal attributes of simultaneous wind instrument timbres: I. von Bismarck's adjectives. *Music Perception, 10*(4), 445 - 468.

Kidd, G. R., & Watson, C. S. (2003). The perceptual dimensionality of environmental sounds. *Noise Control Engineering Journal, 51*(4), 216-231.

Lageat, T., Czellar, S., & Laurent, G. (2003). Engineering hedonic attributes to generate perceptions of luxury: Consumer perception of an everyday sound. *Marketing Letters, 14*(2), 97-109.

Lyon, R. H. (2000). *Designing for product sound quality.* New York: Marcel Dekker, Inc.

Marcell, M. E., Borella, D., Greene, M., Kerr, E., & Rogers, S. (2000). Confrontation naming of environmental sounds. *Journal of Clinical and Experimental Neuropsychology, 22*(6), 830-864.

McAdams, S. (1993). Recognition of sound sources and events. In A. S. Adams & E. Bigand (Eds.), *Thinking in sound: The cognitive psychology of human audition* (pp. 146 -198). New York: Oxford University Press.

McCauley, C., Parmelee, C. M., Sperber, R. D., & Carr, T. H. (1980). Early extraction of meaning from pictures and its relation to conscious identification. *Journal of Experimental Psychology: Human Perception and Performance, 6*(2), 265-276.

Medin, D. L., Lynch, E. B., & Solomon, K. O. (2000). Are there kinds of concepts? *Annual Review of Psychology, 51*, 121-147.

Medin, D. L., & Barsalou, L. W. (1987). Categorization process and categorical perception. In S. Harnad (Ed.), *Categorical perception: The groundwork of perception*. Cambridge: Cambridge University Press.

Mervis, C. B., & Rosch, L. (1981). Categorization of natural objects. *Annual Review of Psychology, 32*, 89-115.

Murphy, G. L. (2002). *The big book of concepts*. Cambridge, MA: MIT Press.

Murphy, G. L., & Smith, E. E. (1982). Basic-level superiority in picture categorization. *Journal of Verbal Learning and Verbal Behavior, 21*, 1-20.

Orgs, G., Lange, K., Dombrowski, J. H., & Heil, M. (2006). Conceptual priming for environmental sounds and words: An ERP study. *Brain and Cognition*, *62(*3), 267-272.

Ortony, A., Clore, G. L., & Collins, A. (1988). *The cognitive structure of emotions*. Cambridge: Cambridge University Press.

Özcan, E., & van Egmond, R. (2005). Characterizing descriptions of product sounds*. Proceedings of the International Conference on Auditory Display, Ireland, 11,* 55-60*.*

Özcan, E., & van Egmond, R. (2006). *Product sound design and application: An overview. Proceedings of the International Conference on Design and Emotion, Sweden*, *5,* 1-19.

Özcan, E., & van Egmond, R. (2007). Memory for product sounds: The effect of sound and label type. *Acta Psychologica, 126*(3), 196-215.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology - Revue Canadienne De Psychologie, 45*(3), 255-287.

Petocz, A., Keller, P. E., & Stevens, C. (in press). Auditory warnings, signal-referent relations, and natural indicators: Re-thinking theory and application. *Journal of Experimental Psychology: Applied.*

Porteous, D., & Mastin, J. F. (1985). Soundscape. *Journal of Architectural Planning Research, 2*(3), 169-186.

Palmer, S. E. (1975). The effects of contextual scenes on the identification of objects. *Memory & Cognition, 3*(5), 519-526.

Pecher, D., Zeelenberg, R., & Barsalou, L. W. (2003). Verifying different-modality properties for concepts produces switching costs. *Psychological Science, 14*(2), 119-124.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (1995). Random Numbers. In *Numerical recipes in C: The art of scientific computing* (2nd ed., pp. 274-316). Cambridge, UK: Cambridge Press.

Rosch, E. H. (1978). Principles of categorization. In E. H. Rosch & B. B. Lloyd (Eds.), *Cognition and categorization* (pp. 27-48). Hillsdale, NJ: Lawrence Erlbaum.

Saygin, A. P., Dick, F., & Bates, E. (2001). Linguistic and nonlinguistic auditory processing in aphasia. *Brain and Language, 79*(1), 143-145.

Solomon, K. O., Medin, D. L., & Lynch, E. (1999). Concepts do more than categorize. *Trends in Cognitive Sciences, 3*(3), 99 - 105.

Solomon, L. N. (1958). Semantic approach to the perception of complex sounds. *Journal of the Acoustical Society of America, 30*(5), 421-425.

Spence, C., & Zampini, M. (2006). Auditory contributions to multisensory product perception. *Acta Acustica, 92*(6), 1009-1025.

Susini, P., McAdams, S., Winsberg, S., Perry, I., Viellard, S., & Rodet, X. (2004). Characterizing the sound quality of air-conditioning noise. *Applied Acoustics, 65*(8), 763-790.

Te Linde, J. (1983). Pictures and words in semantic decisions. In J. C. Yuille (Ed.), *Imagery, memory and cognition* (pp. 117-130). New Jersey: Lawrence Erlbaum.

Thompson, V. A., & Paivio, A. (1994). Memory for pictures and sounds: independence of auditory and visual codes. *Canadian Journal of Experimental Psychology, 48*(3), 380-398.

Tversky, B., & Hemenway, K. (1984). Objects, parts, and categories. *Journal of Experimental Psychology: General, 113*(2), 169-197.

Van Egmond, R. (2004). Emotional experience of frequency modulated sounds: implications for the design of alarm sounds. In D. de Waard, K. A. Brookhuis & C. M. Weikert (Eds.), *Human factors in design* (pp. 345-356). Maastricht: Shaker Publishing.

Van Egmond, R. (2008). The experience of product sounds. In H. N. J. Schifferstein & P. Hekkert (Eds.), *Product experience*. Amsterdam: Elsevier.

Van Rompay, T. (2008). Product expression: Bridging the gap between the symbolic and the concrete. In H. N. J. Schifferstein & P. Hekkert (Eds.), *Product experience*. Amsterdam: Elsevier.

Van Rompay, T., Hekkert, P., Saakes, D., & Russo, B. (2005). Grounding abstract object characteristics in embodied interactions. *Acta Psychologica, 119*(3), 315-351.

Vanderveer, N. J. (1979). Confusion errors in identification of environmental sounds. *Journal of the Acoustical Society of America, 65*, S60-S60.

Västfjäll, D., Gulbol, M. A., Kleiner, M., & Gärling, T. (2002). Affective evaluations of and reactions to exterior and interior vehicle auditory quality. *Journal of Sound and Vibration, 255*(3), 501-518.

Västfjäll, D., Kleiner, M., & Gärling, T. (2003). Affective reactions to interior aircraft sounds. *Acta Acustica, 89*(4), 693-701.

Västfjäll, D., Kleiner, M., & Gärling, T. (2003). Affective reactions to and preference for combinations of interior aircraft sound and vibration. *International Journal of Aviation Psychology, 13*(1), 33-47.

Von Bismarck, G. (1974). Timbre of steady sounds: A factorial investigation of its verbal Attributes. *Acustica, 30*, 146 - 159.

Warren, R. M. (1993). Perception of acoustic sequences: Global integration versus temporal resolution. In A. S. Adams & E. Bigand (Eds.), *Thinking in sound: The cognitive psychology of human audition* (pp. 37-68). New York: Oxford University Press.

Yost, W. A. (1991). Auditory image perception and analysis: The basis for hearing. *Hearing Research, 56*(1-2), 8-18.

Zampini, M., Guest, S., & Spence, C. (2003). The role of auditory cues in modulating the perception of electric toothbrushes. *Journal of Dental Research, 82*(11), 929-932.

Zar, J. H. (1996). Measures of dispersion and variability. In *Biostatistical analysis*. New Jersey: Prentice Hall.

Zwicker, E., & Fastl, H. (1990). *Psychoacoustics: Facts and models*. Berlin: Springer.

This chapter is based on the paper:

Özcan, E., & van Egmond, R. (2005). Characterizing descriptions of product sounds*. Proceedings of the International Conference on Auditory Display, Ireland, 11,* 55-60.

Abstract

Listeners use different types of descriptions for domestic product sounds depending on the level of identification. By conducting labeling and identification tasks, we classified these descriptions into 11 semantically different groups. These groups are organized within a perceptual framework that describes the identification process of product sounds. The results of this investigation indicate that product sounds have associated meanings. This study not only provides an insight into how people perceive and identify product sounds but also supplies preliminary structured information in order to create an exclusive lexicon for product sounds.

CHAPTER 2

⁂

CHARACTERIZING DESCRIPTIONS OF PRODUCT SOUNDS

It is remarkable to see the diversity of associative meanings a single product sound may convey. Our research showed that listeners described a hair dryer sound as air, blowing, soft, relaxed, bathroom, a small vacuum cleaner, and inevitably as a hair dryer. Each one of these descriptions certainly transmits various aspects of the perceived product sound. However, if the presented sound consists of the same spectral-temporal composition, then, how and why do listeners come up with different types of descriptions?

The diversity of meanings may be a result from the extent to which people are able to encode the ‘sound information’. This encoding may be on an acoustical, psychoacoustical, perceptual, cognitive, or emotional level. The descriptions will then be dependent on the level at which a sound is encoded. For example, perceptual (e.g., timbre), cognitive (e.g., mental representation, context, environment), and acoustic (e.g., frequency distribution of the sound) assessments of everyday sounds may be related to the identification process and the categorization for such sounds (Ballas, 1993). If the product sound domain is considered as the sub-domain of everyday sounds, Ballas’ findings serve as a starting point for understanding how listeners treat product sounds. Moreover, examining the type of descriptions people give for product sounds will provide insight into what people hear and how they interpret product sounds. In other words, semantic associations of product sounds may possibly be obtained by analyzing the diverse sound descriptions.

In the field of auditory perception, there have been various investigations on the semantics of everyday and man-made sounds. Solomon (1958) investigated the descriptive adjectives that characterize passive sonar sounds. Von Bismarck's (1974) investigation on timbre of steady sounds revealed that the attribute sharpness was the salient factor, which also represented the other adjectives like hard, loud, angular, tense, unpleasant, bright, high, obtrusive. Björk (1985) showed that natural sounds have emotional associations, and five dimensions (evaluation: pressing, tense, unpleasant; etc.; activity: dull, mellow, hazy, etc.; potency: powerful, loud; simplicity: simple, patterned; and fast-slow scale: fast) were sufficient to describe such sounds. Edworthy, Hellier, and Hards, (1995) investigated the potential meanings associated with warning signals showing that the changes in acoustic dimensions (pitch, speed, inharmonicity, and rhythm) affect the meaning, and consequently, the adjectives (controlled, dangerous, steady, urgent, etc.) that describe warning sounds. Bonebright (2001) showed that acoustical and psycho-acoustical attributes of everyday sounds constitute the perceptual structure for such sounds. These studies suggest that everyday sounds are represented in a listener's mind in various ways, e.g., as encoded acoustical information, emotional experiences, structural properties, and, therefore, have associated meanings. However, these findings still do not explain why some descriptions are more frequent. In this sense, analyzing the identification process of product sounds is likely to provide insight into which aspects of product sounds listeners perceive. We conducted a series of experiments to study how listeners categorize and label product sounds. The findings were gathered within a perceptual framework that describes the identification process of product sounds (see Chapter 1). According to this framework, three main consecutive stages (i.e., perception, recognition, and identification) constitute the product sound identification process resulting in three levels of outputs: descriptions of structural, emotional, and acoustical properties (no recognition); location and/or action description (recognition with loose associations); sound source description (perfect identification).

In this paper, the functioning of three main stages and the reason for three different levels of outputs is discussed. In addition, there is a detailed explanation of product sound specific vocabulary with respect to the levels of identification.

Identification process

Psychologists in different fields have been interested in the processes by which people perceive objects. In the field of visual perception, Biederman's (1981) recognition by components theory explains the visual object categorization process. The mnemonic theory of odor perception combines the odor information processing and its implication for cognitive functions such as recognition, learning, priming, memory, and imagery (Stevenson & Boakes, 2003). In the field of auditory perception, Bregman (1990) proposes two important concepts that underlie the mental process of how people come to perceive events (acoustical sequences): auditory stream segregation, and/or auditory stream integration. According to McAdams' model (1993), auditory objects/events pass through the stages of auditory processing: sensory transduction, auditory grouping, analysis of auditory properties and/or features, and lexicon matching. These studies indicate that several stages are involved in the course of categorization and recognition processes, but do not provide enough evidence of the relationship between the stages and the description of the perceived stimuli at those stages. As sound descriptions can be used as shortcuts to retrieve meaning from memory (Bartlett, 1977; Chiu & Schacter, 1995) and they represent the perceptual qualities of a sound, analyzing descriptions given for a product sound will probably help to understand the consecutive steps in the identification process of product sounds.

In an earlier study, the identification process of product sounds has been presented in the form of a framework (see Chapter 1). In this process, continuous sensorial assessment and information exchange exist within high-level perceptual and cognitive functions. Resulting is a matching mental representation for the perceived auditory stimulus expressed as sound descriptions. The identification process is triggered by any sound generated by the interaction of different parts and materials of a product. Engines, fans, gears, doors, any kind of switches/buttons, flowing liquids, digital devices generate sounds that exemplify some of the product sounds. First, frequency content and temporal characteristics of the sound are analyzed with respect to, e.g., the loudness of the sound. Once spectral and temporal analysis is completed, perception occurs upon the sensory experience. Next, the perceived product sound is matched with the mental representation of the sound in the auditory memory with similar properties. So far the identification process makes use of Bregman's (1990) auditory stream segregation, and/or auditory stream integration concepts and also resembles McAdams' (1993) auditory identification model. However, this framework tries to elaborate on the identification process in a way that the degree of identification determines the type of description used for the product sound event.

Studies indicate that listeners are able to perceive the material, size, and shape of sounding objects and describe the perceptual qualities of sounds (Hermes, 1998; Kunkler-Peck & Turvey, 2000; Klatzky, Pai, & Krotkov, 2000). They can also extract acoustical information from, and structural and temporal properties of, a sound, experience basic emotions upon hearing a sound, and access prior knowledge related to the location in which they heard the sound (Ballas, 1994; Björk, 1985; Bregman, 1990; Edworthy et al., 1995; Gaver, 1993; Gygi, Kidd, & Watson, 2004; Solomon, 1958; van Egmond, 2004; von Bismarck, 1974). This framework therefore aims to organize all these possible perceptual qualities of product sounds and to specify at which stage of the identification process certain perceptual qualities of a product sound occur. If there is no match between the perceived product sound and any mental representation, subsequently, no recognition takes place. Listeners, at this stage, tend to describe the perceived sounds using more high-level concepts because the only available information is the structural and acoustical properties of sound (e.g., droning, continuous, sharp, high-pitched). Because of the lack of information to associate the sound with any meaning in memory, listeners tend to experience positive or negative basic emotions, e.g., like-dislike; pleasant-unpleasant. Damasio (1999) also explains that emotions provide an immediate response to challenges that a person is faced with. If the properties of the perceived product sound match a mental representation in memory, recognition is then completed.

After recognizing a sound, a listener attempts to identify the source of the sound. In this framework, the aim of an identification process is defined as labeling the product sound with the sound source identification (e.g., fan, engine, hairdryer). However, the other types of sound descriptions at previous stages are also considered as sound identification that varies in degrees of association. If the sound source is recognized but cannot be identified, listeners describe the sounds by location of the sound and the action that generates the sound (e.g., something rotating, bathroom, house). Gaver's (1993) map of everyday sounds supports this assumption suggesting that sound conveys information about events at locations in an environment.

The process described for identifying a product sound is apparently based on the knowledge from other studies in the field auditory of perception. Several labeling and identification studies have been conducted to support this framework in terms of product sound perception (see Chapter 1), in the next section, the qualitative analysis regarding the product sound descriptions obtained by these studies is discussed.

Product sound descriptions

In the labeling and identification experiments, participants labeled the groups of perceptually similar sounds (free labeling) and finally tried to identify individual product sounds by using short written descriptions which they thought may describe the presented product sound (free identification) (see Chapter 1). All these labels/sound descriptions were collected to provide insight into different levels of associations that a product sound may have.

Sound labels

Free labeling and identification tasks yielded different product sound descriptions. Strong similarities among the sound descriptions were observed; so the obtained descriptions needed to be classified. The classification of the descriptions was performed as follows: first, each word typed by a subject in a free labeling and an identification task was extracted from its meaningful combination of words; resulting was a list of 1000 single words (e.g., ‘big washing-machine in distance’ as ‘big’, ‘washing-machine’, ‘distance’). Second, all these words were analyzed to see if there were any conspicuous patterns in the way listeners describe product sounds. Based on the found patterns and similarities within patterns, the descriptions were classified into 11 groups: action, emotion, location, material, meaning, onomatopoeia, psycho-acoustical properties, sound type, source, source properties, and temporal aspects of sound. Finally, each single extracted word was scored as ‘1’ if the word corresponded to any of the pre-defined description groups. For example, the words ‘big’, ‘washing-machine’, and ‘distance’ were rated as ‘1’ respectively in the source properties, source, and location groups.

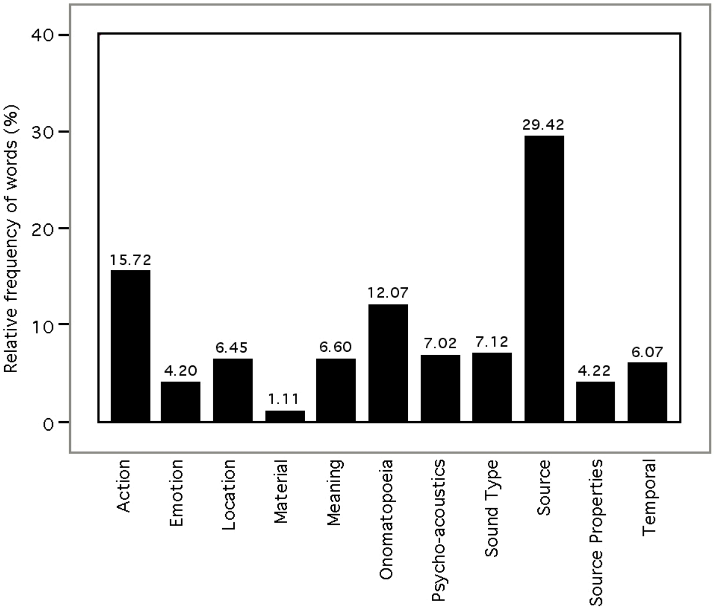


Figure 1. Relative frequency of words as a function of product sound descriptive groups. (The cumulative percentage over descriptive groups add up to 100%).

Figure 1 presents the relative frequency of words as a function of product sound descriptive groups. Of all the descriptions 29.42% was used for source descriptions and 15.72% for action descriptions indicating the listeners' main concern in describing product sounds. The descriptive groups are further explained in the following paragraphs.

The sound *source labels* constitute the main bulk of the product sound descriptions. These labels describe the main product in operation (e.g., microwave oven), the special part of the product which causes a secondary sound (e.g., beep, door, or rotary buttons on a microwave oven), and the medium which the sound is produced in (e.g., air, liquid, water). Listeners also describe additional information about the source, *source properties*. Most of these descriptions are adjectives: 'cold' water, 'heavy' door, and 'old' typewriter. The high percentage for the source group shows the listeners' tendency to identify a product sound with a source.

The *action* descriptions constitute the second essential product sound descriptive groups. These descriptions contain various verb phrases describing the action that causes the sound without an actor, for example, opening or closing door, pouring water, finishing beep, turning button, hitting with an object. This indicates that a listener can identify both the source (i.e., a specific part of the product) and the action causing the sound. However, some action descriptions do not take objects with the verb. Some examples are 'blowing, moving, droning, operating, running, and cleaning'. This indicates that a listener can hear the action but does not specify the source of the sound (e.g., washing machine, hair dryer). The latter descriptions exhibit a general idea about the action of the perceived sound. All these descriptions also indicate that the product operation cycle results in sounds specific to the movement of the working parts in a product.

The descriptions in the *onomatopoeia* group occur when a listener is unable to extract enough information to identify the perceived product sound. If the sound evokes no meaning, the listener simply tries to imitate the product sound by generating similar sounds to the original sound (e.g., brrr, kling). The analogy would be sketching the main descriptive features of a strange object. Another version of this type of sound description is using the conventional vocabulary for imitating sounds such as 'buzzing, rattle, beep, droning, plop, hum'. These conventional words are also used as verbs which describe action, indicating that listeners are able to associate the sounds with certain actions that cause sound (e.g., clicking with a mouse button).

The descriptions in the *psycho-acoustical* properties group describe the acoustical and psycho-acoustical aspects of the product sounds. Some examples are 'soft, amplified, monotone, high-pitched, low-pitched, sharp, smooth, round, quiet, edgy, loud'.

All these sound descriptive groups presented suggest that product sounds convey various meanings with relation to the location, action, and source identification. On the other hand, product sounds also convey meanings that are not linked to physically identified features, but linked to rather *abstract meanings*. For example, listeners often interpret alarm clock sounds as 'wake-up, attention' or 'time is up!', microwave oven sound as '10 seconds' or 'the food is ready', warning signals and malfunctioning products as 'danger, error, emergency, activate, malfunction'. In some cases, listeners use possession descriptions for product sounds, such as 'my alarm clock'.

The descriptions in the *sound type* group specify the means by which a sound is produced. These include 'digital, mechanical, electrical, electronic, analogue, metallic, synthetic, aerodynamic'.

The descriptions in the *location* group contain contextual information where a product sound may take place, e.g., bathroom, laundry, house, outside, hospital, big room, public, domestic'. These labels indicate that product sounds have contextual associations and certain locations have their own soundscape related to product sounds.

The descriptive words concerning the *temporal* aspects of product sounds indicate that listeners pay attention to the temporal information that product sounds convey. Temporal descriptions include 'short, long, continuous, repetitive, multiple, single, and constant'.

The descriptive words in *emotion* group indicate that listeners have emotional experiences upon hearing a product sound. Most of the descriptions indicate the acceptability of the product sounds and evoke basic emotions such as like-dislike; they include 'annoying, not annoying, unbearable, relaxing, boring, acceptable, angry, happy, warm, cozy, irritating, disturbing'. As sound is as a consequence of an operating product, most of the time listeners cannot intervene the occurred sound. This sometimes causes basic negative emotions, such as 'irritating' or 'disturbing'.

The descriptive words in *material* group mostly describe the material component of two interacting objects. These descriptions are also very closely related to the source of the sound and source properties. They simply describe the materials in interaction, e.g., plastic, metal, and wooden.

Product sound groups

In an earlier study, we classified product sounds into six groups: air, alarm, cyclic, impact, liquid, and mechanical sounds (see Chapter 1). The same sounds and sound groups were used for the free labeling and free identification task. It was observed that the descriptions vary from one group of product sounds to another and each product sound group has characterizing descriptions. We determined the frequency count for each descriptive category in combination with the sound groups. Thus, a 6 by 11 frequency matrix resulted. Correspondence analysis was used to analyze this frequency data in order to reveal the association between descriptive groups and sound groups. A 3-dimensional solution explained 96% of the variance (Figure 2).

Figure 2 illustrates how sound descriptive groups are distributed over product sound groups in 3-dimensional space. According to the figure a product sound group is well described by the closest sound description around it:

- Air sounds by location, action and psycho-acoustical description of sounds
- Alarm sounds by abstract meanings which the sound conveys
- Cyclic sounds by location, sound type and psycho-acoustical descriptions of sounds
- Liquid sounds by the action causing the sound and sound source
- Impact sounds by the temporal aspects of the sound, the special properties of the sound source, onomatopoeias, and interacting materials
- Mechanical sounds by emotional experiences.

The distribution of descriptive words over product sound groups showed that 33.96% was accounted for impact sounds, followed by 15.69% for air sounds, 13.10% for cyclic sounds, 12.59% for alarm. 12.44% for mechanical sounds, and 12.22% for liquid sounds. This may indicate the difficulty of identifying the source of impact sounds resulting in listeners using diverse descriptions for impact sounds to compensate the difficulty. On the other hand, probably because the sources of the liquid sounds are well identified, listeners do not need to elaborate on the descriptions for such sounds.

Figure 3 presents the relative frequency of product sound descriptive groups as a function of product sound groups. According to the figure, listeners describe product sounds mostly by sources independently of the sound type. Alarm sounds are mostly described by the meanings that they convey. Onomatopoeias are mostly used for alarm, impact, and mechanical sounds. The product sound groups with their characterizing descriptions are further explained as follows:

The *air* sounds are mostly described by source (31.42%), action (16.53%), sound type (12.84%), psycho-acoustical properties (11.88%), and location descriptions (9.84%). Some examples are ‘vacuum cleaner, hair dryer, motor, drying, vacuuming, aerodynamic, sharp, loud, noise, and industry’.

The *alarm* sounds are mostly described by the conveyed meanings (25.89%), source descriptions (22.32%), and onomatopoeias (13.97%). Some examples are ‘bell, microwave oven, beep, buttons on a microwave, warning, alarm, attention’. Not many action descriptions are used for alarm sounds, perhaps because there is no visible action causing sound. Furthermore, few location descriptions are used, perhaps because alarm sounds are supposed to warn listeners in and out of contextual situations.

The *cyclic* sounds resemble air sounds. They are mostly described by source (27.99%), location (14.89%), psycho-acoustical properties (13.09%), and action descriptions (12.60%). Some examples are ‘vacuuming, blowing, dryer, fan, monotone, soft, low pitch, laundry room, ventilator, background’. Few material source property descriptions are used to describe such sounds.

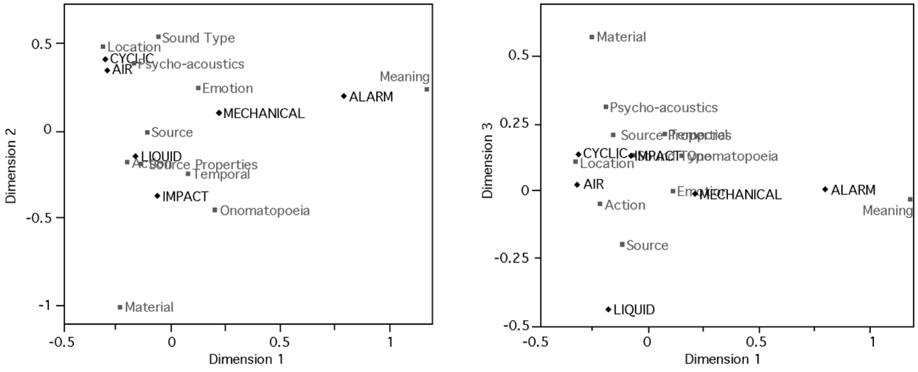


Figure 2. *A 3-dimensional correspondence analysis solution of the frequency data of descriptive groups and sound groups. Upper graph shows dimension 1 vs. 2, and the lower graph dimension 1 vs. 3. The text in upper case represents product sound groups, the text in lower case represents the product sound descriptive groups.*

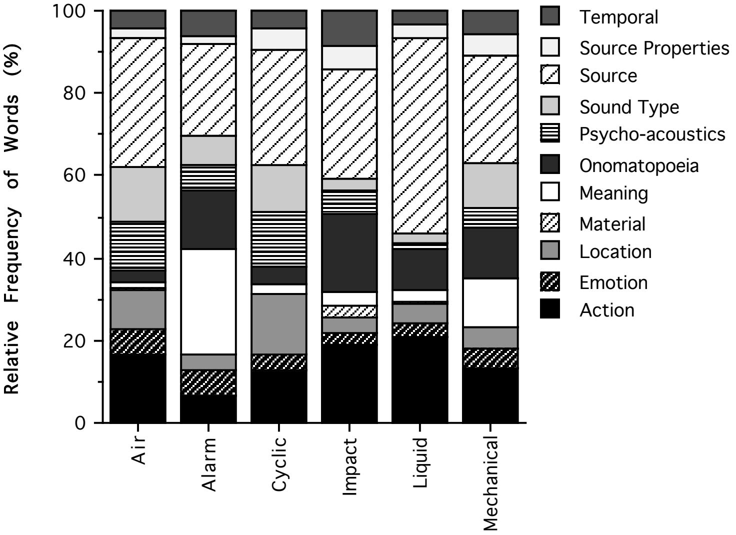


Figure 3. *Relative frequency of words as a function of product sound descriptive groups and of product sound groups.*

The *impact* sounds are mostly described by source descriptions (26.33%), action descriptions and onomatopoeias (19%), and temporal aspects of impact sounds (8.59%). Some examples are ‘door, switch, short, single, click, bang, opening’. Few emotion and location descriptions are used for these sounds.

The *liquid* sounds are mostly described by source (47.37%), action descriptions (20.70%), and onomatopoeias (10%) Some examples are ‘coffee machine, boiling, home, pouring, filling, bubble’. Few psycho-acoustical descriptions are used to describe such sounds.

The *mechanical* sounds are mostly described by sources (26.38%), action descriptions (13.45%), onomatopoeias (12.59%), abstract meanings (11.55%), and sound types (10.86%). Some examples are ‘adjusting, rotating, rattling, shaver, buzzer, danger, high-pitch, electro-motor, mechanical’. Listeners often describe these sounds by source and action description, but not by location or material description of the sound.

Only the salient descriptions that characterize each product sound group have been presented above. However, product sounds exhibit more associated meanings than the presented above. Upon the perception of a product sound, any associated meaning can be retrieved from memory for that sound, i.e., a product sound can evoke meanings on all levels of identification emerging in any type of descriptions.

For example, the results also show that impact sounds material descriptions (90.38%), psychoacoustics (60.67), onomatopoeias (53.46%), temporal descriptions (48.06%), source properties (46.11%), and action descriptions (41.06%) are mostly used for impact sounds; abstract meanings for alarm sounds (49.35%); psychoacoustics for cyclic (45.66%). Emotion and source descriptions are distributed over all sound groups, indicating that any product sound can evoke positive or negative emotional experiences.

Conclusions

The main purpose of this study was to discover whether product sounds have associations with meanings on different levels of a product sound identification process (perceiving, recognizing, and identifying) and if so, to reveal the types of descriptions that characterize product sounds on these different levels. The results suggest that the sound descriptions given by the participants can be organized into 11 groups. These descriptions are based on the levels of the previously presented identification process for product sounds according to their identifiablity degree.

A product sound cannot be recognized or identified, if a matching meaningful association does not exist in the long-term memory. In such a situation, listeners can only verbalize their percept of the product sound while the perceived information is still in their working memory. According to Vanderveer (1979), listeners describe the sensory qualities of sounds in the case of not identifying the source of an object. So, onomatopoeias, psycho-acoustical and temporal descriptions of the product sounds, and emotional experiences are on this level. The descriptions here are very vague and can be used for any product sound. At the recognition level, a product sound is recognized with good match but loosely associated with the mental representation. Listeners can retrieve some information related to the perceived product sound, but prior knowledge is necessary at this level. The action that causes the sound, the location where a product sound is frequently heard, the interacting material descriptions, and sound types are organized in this level. Here the descriptions are more specific to a limited number of product sounds. At the identification level, a product sound is identified with good association, and the product labels or labels for the secondary parts of the product that causes the sound are used. Source property descriptions also appear at this level. The abstract meanings that a listener derives from a product sound are also categorized in this level. Here, listeners retrieve the exact information to describe a product sound. These descriptions are very specific to the source or the meaning of the product sounds.

Identifying a product sound with good association (i.e. perfect identification) allows listeners to retrieve all other meanings that a product sound may convey. Thus, describing a product sound as ‘shaver’ activates all other meanings that a shaver sound may be associated with, such as, high-pitch, continuous, annoying at no-recognition level; shaving, bathroom at recognition level; and shaver at identification level. It can therefore be concluded that product sounds have associated meanings.

The variety of words used to describe the source of a product is another indication for the objective of the identification process. Because sound is a consequence of objects in interaction, listeners tend to first identify the product sounds first by the source, then by the action that causes the sound. However, not all the chosen words exhibited correct identification with good association with the original product. Perfect identification occurs only when a listener is able to extract the source information from the perceived product sound. Whether the listener reaches the exact association in terms of source description is not of interest to this investigation. Because, while a listener may identify a product sound, e.g., as hair dryer instead of vacuum cleaner, as the listener comes up with a source description but not an action or location description, in the identification process this imperfect identification makes no difference. In addition, when listeners find it difficult to recognize a sound source with good match, they tend to use the combinations of many words to describe a product sound to compensate for their failure in identifying the exact source of the sound.

The results show that material descriptions constitute only about 1% of the all the given descriptions. This may seem contradictory to Gaver’s (1993) framework of everyday sound listening. Gaver, in his framework, classified the sound sources of everyday events into three groups of interacting materials: vibrating objects, aerodynamic sounds, and liquid sounds. In our framework, air and liquid descriptions are classified in the source descriptions, not particularly in material descriptions, and source descriptions constitute the main bulk of product sound descriptions. Moreover, from a product design point of view, a product can be expressed by texture, color, shape, and materials choice. Materials indicate what a product is made of, e.g., plastic, wood, metal, etc. In this sense, the results show that 90% of the material descriptions were given for impacting product sounds. Thus, suggesting that listeners can hear the material of a product when the product or part of it is in interaction with other parts, e.g., switches, rotary buttons, or doors. In addition, Gaver’s suggested framework describes everyday sounds relying on an ecological account, whereas our conclusions are based on empirical findings and concern only product sounds.

Discussion

This study has provided insight into how people perceive and identify product sounds. In addition, the analysis of the verbal attributes revealed 11 headwords that supply preliminary structured information to create a lexicon for product sound. Such a lexicon can be used in different fields such as product sound perception, design, and application. In line with the auditory display design, one of the main purposes of our investigations on product sounds is to develop a computer-based tool by which designers will be able to model the conceptual ideas for the sound design of products. Thus, these sound descriptions will indicate what aspects of product sounds listeners perceive and how these perceived aspects of sounds can be modified and manipulated in order to create an acceptable sound for the products.

References

Ballas, J. A. (1993). Common factors in the identification of an assortment of brief everyday sounds. *Journal of Experimental Psychology: Human Perception and Performance, 19*(2), 250-267.

Ballas, J. A. (1994) Effect of event variations and sound duration on identification of everyday sounds*.* *Proceedings of the 2nd International Conference on Auditory Display*, *USA, 2*, 147-155.

Bartlett, J. C. (1977). Remembering environmental sounds: Role of verbalization at input. *Memory & Cognition, 5*(4), 404-414.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review, 94*(2), 115-147.

Bismarck, G. (1974). Timbre of steady sounds: A factorial investigation of its verbal attributes. *Acustica, 30*, 146 159.

Björk, E. A. (1985). The Perceived Quality of Natural Sounds. *Acustica, 57*(3), 185-188.

Bonebright, T. L (2001) *Perceptual structure of everyday sounds: A multidimensional scaling approach*. Proceedings of the *International Conference on Auditory Display, Finland, 7*, 73-78.

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound.* Massachusetts: MIT Press.

Chiu, C. Y. P., & Schacter, D. L. (1995). Auditory priming for nonverbal information: Implicit and explicit memory for environmental sounds. *Consciousness and Cognition, 4*(4), 440-458.

Damasio, A. (1999). *The feeling of what happens: Body, emotion and the making of consciousness*: London: Vintage.

Edworthy, J., Hellier, E., & Hards, R. (1995). The semantic associations of acoustic parameters commonly used in the design of auditory information and warming signals. *Ergonomics, 38*(11), 2341-2361.

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

Gygi, B., Kidd, G. R., & Watson, C. S. (2004). Spectral-temporal factors in the identification of environmental sounds. *Journal of the Acoustical Society of America, 115*(3), 1252-1265.

Hermes, D. J. (1998). *Auditory material perception* (Annual Progress Report No. 33): IPO.

Klatzky, R. L., & Pai, D. K., & Krotkov, E. (2000) Perception of materials from contact sounds,. *Presence* 9(4), 399-410, 2000.

Kunkler-Peck, A. J., & Turvey, M. T. (2000). Hearing shape. *Journal of Experimental Psychology: Human Perception and Performance, 26*(1), 279 - 294.

McAdams, S. (1993). Recognition of sound sources and events. In A. S. Adams & E. Bigand (Eds.), *Thinking in sound: The cognitive psychology of human audition* (pp. 146-198). New York: Oxford University Press.

Özcan, E. & van Egmond, R., & Jacobs, J. (submitted) Bases for categorization and identification of product sounds.

Solomon, L. N. (1958). Semantic approach to the perception of complex sounds. *Journal of the Acoustical Society of America, 30*(5), 421-425.

Stevenson, R. J., & Boakes, R. A. (2003). A mnemonic theory of odor perception. *Psychological Review, 110*(2), 340-364.

Van Egmond, R. (2004). Emotional experience of frequency modulated sounds: implications for the design of alarm sounds. In D. de Waard, K. A. Brookhuis & C. M. Weikert (Eds.), *Human factors in design* (pp. 345-356). Maastricht: Shaker Publishing.

Vanderveer, N. J. (1979). Confusion errors in identification of environmental sounds. *Journal of the Acoustical Society of America, 65*, S60-S60.

This chapter is based on the paper:

Özcan, E., & van Egmond, R. (2007). Memory for product sounds: The effect of sound and label type. *Acta Psychologica*, *126*(3), 196-215.

Abstract

The (mnemonic) interactions between auditory, visual, and the semantic systems have been investigated using structurally complex auditory stimuli (i.e. product sounds). Six types of product sounds (air, alarm, cyclic, impact, liquid, mechanical) that vary in spectral-temporal structure were presented in four label type conditions: self-generated text, text, image, and pictogram. A memory paradigm that incorporated free recall, recognition, and matching tasks was employed. The results for the sound type suggest that the amount of spectral-temporal structure in a sound can be indicative for memory performance. Findings related to label type suggest that 'self' creates a strong bias for the retrieval and the recognition of sounds that were self-labeled; the density and the complexity of the visual information (i.e., pictograms) hinders the memory performance ('visual' overshadowing effect); and image labeling has an additive effect on the recall and matching tasks (dual coding). Thus, the findings suggest that memory performance for product sounds are task-dependent.

CHAPTER 3

⁂

MEMORY FOR PRODUCT SOUNDS: THE EFFECT OF SOUND AND LABEL TYPE

During the last few decades, connections between various perceptual systems (visual, auditory) and the verbal semantic system (Paivio, Philipchalk, & Rowe, 1975) and the extent to which inter-system relationships (verbal vs. visual) influence memory performance have been extensively studied (Paivio, 1991; Thompson & Paivio, 1994; Schooler & Engstler-Schooler, 1990; Melcher & Schooler, 1996). These studies suggest that object information is organized into perceptual and semantic systems within the cognitive system. On the one hand, it has been shown that semantic (i.e., verbal) information congruent with perceptual (i.e., visual) information has an additive effect on memory (i.e., *dual-coding*, Paivio, 1991). On the other hand, it has been found that verbalization of the semantic information acquired from a complex stimulus might impede recognition (i.e., *verbal overshadowing*, Schooler & Engstler-Schooler, 1990). These theories imply that the interaction between a perceptual system and a semantic system could be mnemonic to a certain extent. Similarly, with this study we investigate how the memory for the auditory system is affected by the added presence of semantic information. For this, we focus on product sounds—a sub-category of environmental sounds—that vary in spectral-temporal structure. Using a memory paradigm, we investigate the effect of visual and verbal information on the encoding of, and memory for, product sounds. In general, the results aim to give insight into mnemonic interactions between multiple perceptual systems and the semantic system.

Product sounds

We define product* sounds as a sub-category of environmental sounds that are emitted by domestic appliances (e.g., vacuum cleaners, dishwashers, alarm clocks, coffee machines). All these appliances have mechanical and electrical parts that produce sounds. The produced sounds have been classified in six categories which have been based on perceptual similarities (see Chapter 1). These categories include: (*i*) *air* sounds (caused by moving air due to the rotating fans used to blow or suck air); (*ii*) *alarm* sounds (mostly digitally produced and designed especially to provide feedback and to warn listeners); (*iii*) *cyclic* sounds (caused by rotating parts which result in a cyclic event and a periodicity in the sound); (*iv*) *impact* sounds (caused by a short impact between product parts); (*v*) *liquid* sounds (caused by moving or heating up liquids); and (*vi*) *mechanical* sounds (caused by engines at high rpm and small rotating, rubbing mechanical parts of products).

The sounds in these categories vary in their spectral-temporal structure. The amount of structure in the spectral composition of product sounds decreases in the following way: (a) highly structured, a pure tonal composition—not containing noise (e.g., alarm sounds), (b) medium structured, harmonic bands caused by the periodicity of the engine together with noise (e.g., mechanical, cyclic, and air sounds), and (c) unstructured, only noise or changing spectral structure over time (e.g., liquid and impact sounds). The amount of structure in the temporal composition of product sounds decreases in the following way: (a) highly structured, rhythmic-like pattern (e.g., alarm sounds), (b) medium structured, periodicity caused by the engine (e.g., mechanical, cyclic, and air sounds), and (c) unstructured, no regularity (e.g., liquid and impact).

It has been shown that people seek for systematic organizations such as structural units and hierarchies during the encoding or retrieval process of the information (e.g., see, Deutsch, 1980; Deutsch & Feroe, 1981; Sternberg, 1998). These structural units come as geometric shapes for visual objects (Palmer, 1977; Biederman, 1987; Liu & Cooper, 2001), and tonal and temporal structures for auditory objects (Deutsch, 1972; Deutsch & Feroe, 1981; Garner, 1974; Povel, 1981). In addition, they facilitate recognition depending on how well the structure is extracted from the visual or auditory object. Because memory favors structure and unstructured sequences impose higher memory load (Deutsch, 1980), we can readily predict that memory performance for many product sounds such as impact or air sounds will be more difficult, whereas for alarm and mechanical sounds the performance will be easier.

The aforementioned studies also suggest that information is encoded on different levels of hierarchies, thus creating a hierarchical network. This implies that if needed, tonal and temporal units in a structured product sound could be extracted at any level of the hierarchical organization facilitating memory. However, this may not be the case for unstructured sounds, as they could be partially encoded as a result of loose or weak bonds in the hierarchical organization of the information. Incomplete encoding may create ambiguity in the sound identification, because multiple semantic associations may occur with the mental representation of an unstructured sound. For one item, memory performance tends to worsen if multiple associations exist to choose from.

Studies dealing with structurally complex stimuli have shown that conceptual and perceptual training may reduce perceptual complexity, facilitate recognition, and improve accurate verbalization (see Melcher & Schooler, 2004). For example, Lehrer (1983) has illustrated the expertise of the specially trained wine tasters and their ability to communicate (i.e., verbalize) even the finest details of wine tasting (i.e., complex perceptual stimulus). Moreover, Sweller and Chandler (1991, 1994) describe situations in which conscious thinking and perceptual assessing occur simultaneously and exceed the capacity of working memory within the framework of *cognitive-load* theory. Accordingly, in order to reduce the cognitive load and to increase the learnability, methods that allow dual-mode presentation techniques have been proposed. Tindall-Ford, Chandler, and Sweller (1997) have shown that participants studying instructional materials in multiple modalities (i.e., audio-text and visual diagrams/tables) perform better than those studying in a single modality (visual-only format).

These studies indicate that design teams may benefit from a professional training specialized on how to encode product sounds and how to increase the efficiency in the product sound related communication. They also indicate that despite the structural complexity of the product sounds, designers should be able to improve their perceptual expertise and to learn to code the spectral-temporal structure of a sound in better details in order to capture subtle differences between similar sounds. Moreover, the presence of additional modalities (e.g., verbalization, visualization) at encoding may help to improve memory. Therefore, the present study will focus on the additive effects of perceptual information on the recall and recognition memory for product sounds and on a listener's ability to match the auditory information to the label it was presented with.

Memory tasks and encoding

The conditions in which an object is coded during acquisition influence how well information is stored in memory. Consequently, recognition and recall memory will be dependent on the information that can be extracted and attributed meaning to (Cleary, 2002; also see Schacter, Cooper, & Delaney, 1990). Cleary (2002) distinguishes an identification and a recognition stage. If there is no identification, then at the recognition stage featural aspects will be analyzed. Conversely, if there is identification then attribution of meaning will be involved at the recognition stage. Moreover, 'self' can provide a strong bias towards self-experienced events/objects and self-generated meanings. Greenwald and Banaji (1989) in a series of experiments have shown that recall memory is better for self-generated items rather than provided items. Carmichael, Hogan, and Walters (1932) presented ambiguous visual forms with two different types of labels to two different groups of participants. When participants were asked to reproduce the visual forms in the form of a drawing, their drawings biased the labels presented with the target visual form. Thus, a matching task may be influenced by the extent to which people are able to reproduce the perceived objects.

Moreover, people may exhibit different performances for different memory tasks. For example, Bahrick and Boucher (1968) have shown that recall accuracy of the verbal codes is dissociated from the visual recognition accuracy for the same item. This implies that recall, recognition, and matching tasks operate differently. Procedurally, a *recall* task does not require the actual processing of the information and is therefore dependent on the internal search within the memory. In addition, a *free* recall task depends highly on one's ability to retrieve prior knowledge. *Recognition* requires active processing and involves comparisons with the prior knowledge and depends on how well the prior knowledge is coded. Structural analysis of the object takes place in order to be able to map the actual information to the stored information. A *matching* task is dependent on the retrieval and comparison of at least two different types of stored information (i.e., structural and semantic) that are conceptually related.

Modality effects on memory

Several studies suggest that labeling (visual or verbal) enhances auditory memory. Such an enhancement depends on how the sound is encoded at the acquisition (Cleary, 2002). The recognition performance may improve if perceptual details of environmental sounds are encoded; however verbalization (i.e., naming of the sounds) at encoding may enhance the free recall performance (Bartlett, 1977), or the identification accuracy (Chiu & Schacter, 1995). Chiu and Schacter (1995; see also, Huss & Weaver, 1996) have shown that at encoding a sound is very likely to be coded together with its label, but not vice-versa indicating a one-way mnemonic link from a perceptual to a semantic store. Another study has also indicated a mnemonic link between a perceptual store and a semantic one: Edworthy and Hards (1999) have shown that auditory warnings with verbal labels are better remembered than auditory warnings with image labels; but if participants creates their own text or image labels, performance is the same for both label types (bias for self-generated items).

The study of Greene, Easton, and LaShell (2001) implies cross-modal interactions. That is, stimuli presented in one modality can be substantially identified or recognized in another modality. In another study, Tindall-Ford et al. (1997) have shown that learning of instructional materials, which is considered as cognitively demanding due to the high-intellectual content of the material, improves due to the presence of auditory text together with visual diagrams instead of visual-only format. Moreover, simultaneous presence of visual and auditory information at encoding has an additive effect for recall performance (Thompson & Paivio, 1994). Similarly, Lyman and McDaniel's (1990) study has shown that elaborating odors with information from different modalities (visual and verbal) increases the probability that olfactory information is retrieved. This might be due to the multiplicity of the retrieval paths.

Aforementioned effects might be explained by the dual coding theory. That is, multiple codes have an additive contribution to memory performance (Paivio, 1991; Thompson & Paivio, 1994). One of the reasons for the additivity effect could be that the conceptual information is activated more than twice in the presence of information from verbal and non-verbal systems, which creates a strong path between the systems. Dual coding—as a theory—also tries to explain the mnemonic relations between the verbal and non-verbal systems in the following way: (*i*) pictures are remembered better than text (due to the high imagery elicited by the images), (*ii*) picture and text combinations have an additive effect on memory (high imagery plus a clear label), and (*iii*) activating the conceptual information may suffice for retrieving codes from either verbal or non-verbal stores (Paivio, 1983). Therefore, it is very likely that memory for product sounds also benefits from dual coding if the sounds are coded with additional modalities at the acquisition.

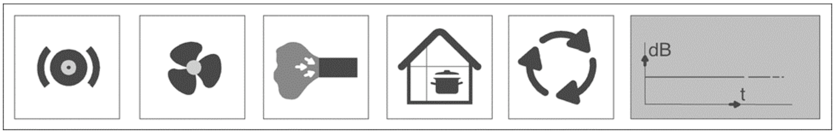
Although dual coding has shown a positive effect of labeling, other studies have shown a reversed effect. In face recognition tests, Schooler and Engstler-Schooler (1990) have coined the term '*verbal overshadowing*' and have demonstrated that if the to-be-remembered non-verbal stimulus is complex and requires fine verbal descriptions, subsequent recognition will be impaired due to verbally biased memory representations. Verbal overshadowing has also been observed for complex auditory stimuli in the form of voice identification (Perfect, Hunt, & Harris. 2002; Vanags, Carrol, & Perfect, 2005). In addition, Melcher and Schooler (1996) have shown both positive and negative effects of verbalization, depending on the expertise: Describing wines from memory produced somewhat of a 'dual-coding' benefit for totally untrained wine drinkers but impaired wine recognition for participants who had some wine training (but were not experts). However, no such effect was found for trained wine experts. In another study, Melcher and Schooler (2004) have suggested that the extent to which recognition memory is disrupted by verbalization depends on the balance between the perceptual expertise and verbal expertise. It seems that a verbal overshadowing effect mostly occurs when perceptual expertise is high but lacks support from verbal expertise. These studies have shown that with conceptual training, verbal expertise can be improved to support the perceptual information. Thus, the link between the two systems may get stronger in favour of recognition performance.

Product sound design related communication

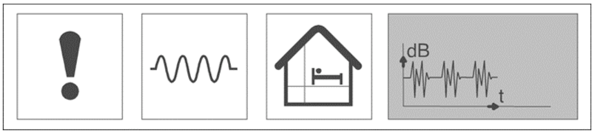
As the field of product sound design is advancing, the need for sound related communication emerges as an essential part of the design teams' regular discussions. Therefore, methods have been proposed in order to support designers' sound related discussions and minimize the load on the cognitive activities (Özcan & van Egmond, 2004 Özcan & van Egmond, 2006; van Egmond, 2006). According to these methods, during the product development designers should 'verbalize' the perceived aspects of the product sound in conversations; 'audiolize' and 'imitate' the concept of the sound by sounding sketches or sounding models; and 'evaluate' the sound quality (i.e., appropriateness of the sound to the product). Basically, these activities involve cognitive functions related to a sound, its corresponding verbal label, and its corresponding visual image (i.e., product itself). Thus, designers are forced to reproduce the sound related information from memory by retrieving the previously stored auditory information which can be on a perceptual, conceptual, or semantic level. For example, if the sound of a vacuum cleaner is in question, the designers may be expected to retrieve the spectral-temporal composition of the sound, the appearance of the product causing the sound and perhaps the product's name, and other conceptual associations that are somehow linked to vacuum cleaner sounds in memory (e.g., similar products, the location, etc.).

Product development in general requires precise information exchange, therefore misunderstandings in the discussions caused by ambiguity should be avoided. Considering the structural complexity of product sounds, verbal communication can be a cognitively challenging task for the design team especially when discussions occur after the exposure to, and in the absence of the sound. Designers are then forced to rely on their limited memory and at the same time they need to make cognitive decisions. As a consequence, communication skills may suffer from ambiguity and poor memory performance.

Moreover, product sound related terminology is often too technical for design teams. Furthermore, members of a design team often have different cultural backgrounds and speak different languages. Martens and Giragama (2002) have shown that the words describing guitar timbres in Japanese and Sinhalese languages (with the same English meanings, e.g., pleasant, cheerful, sharp) are related to different acoustical dimensions. This demonstrates the insufficiency of the verbal communication of



A pictographic depiction of a kitchen hood sound



A pictographic depiction of a digital alarm clock sound

Figure 1. Two pictogram sets are presented to show the type of pictograms used in a set and how the set is formed.

sounds. In order to address these problems, a new method has been suggested that incorporates the visualization of the product sounds in a pictographic manner (Özcan & van Egmond, 2004). The designed pictograms are graphical depictions of sound producing parts, actions that cause sounds, locations where products can be heard, materials, and temporal properties. A combination of these depictions should facilitate the memory for, and, consequently, the communication of product sounds (see Figure 1 for examples). In addition to more conventional sound descriptions, the influence of these pictograms on the auditory memory will be investigated in this study.

Summary and predictions

To see whether labeling would be effective, a memory paradigm that is relevant for product sounds is employed. Similarly to Bartlett's (1977) experimental set up for environmental sounds, three types of memory tasks are incorporated to represent designers memory related cognitive activities. The tasks are *free recall*, *recognition,* and *matching*.

Moreover, four different types of additional semantic information will be used to simulate the main possible encoding methods. The semantic information will be provided in the form of a (*i*) *self labeling* to represent designers' own semantic association (*ii*) *text labeling* to represent the basic semantic association, (*iii*) *picture labeling* to represent directly the object that causes the sound, (*iv*) *pictographic labeling* as an attempt to represent the various perceptual attributes of a sound in one visual composition.

The prediction for the results of this experiment is that visual labels (picture and pictographic) in general will aid the memory for product sounds better than the text labels (self-generated or provided) considering the dual coding theory. Within the text labels (self-generated or given), self-labeling may also be a strong strategy for encoding as studies have shown a bias for the memory for self-labeled objects. Thus, self-generated text labels will outperform the provided text labels. Within the visual labels, picture labeling will have superiority to pictogram labels especially in the recall and recognition tasks considering the cognitive-load theory. With this new concept (pictographic language), we aim to reach the mental representation of a sound by multiple visual information units that in combination refer to the same concept. It is expected that this will help to facilitate the retrieval process better than the other training techniques that only offer one or two retrieval paths with text or images. However, the number of visual information units may also create an 'overshadowing' effect for the retrieval of the information. Moreover, the novelty of the concept of the pictographic representation of product sounds may be a disadvantage for the memory.

To carry out the above-mentioned memory paradigm, six types of product sounds are used that represent the varying ranges of structural complexity of product sounds. We assume that the combination of temporal and spectral structure determines how well a sound can be encoded in memory. Listeners will have the best memory for sounds consisting of a highly structured spectral composition and a highly structured temporal pattern, whereas listeners will have the worst memory for sounds consisting of an unstructured spectral composition and an unstructured temporal pattern. Consequently, recognition and recall scores for product sound categories should decrease in the following way: alarm, mechanical, cyclic, air, liquid, and impact.

Experiment

Method

The experiment was a 4 x 6 mixed factorial design, with label type (self-generated text, text, image, and pictogram) as between-subjects factor and product sound categories (air, alarm, cyclic, impact, liquid, and mechanical) as within-subjects factor. The experiment consisted of four phases: sound-label presentation, free recall, recognition, and matching.

Participants

Seventy-two students (42 male and 30 female) studying industrial design engineering at Delft University of Technology participated. The mean age was 21.3 years. Eighteen students were randomly assigned to each of the four experimental conditions formed by the label type. For the self-generated text label condition, 12 male and 6 female students participated with the mean age of 21.5 years. For the text label condition, 8 male and 10 female students participated with the mean age of 21.9 years. For the image label condition, 10 male and 8 female students participated with the mean age of 20.5 years. For the pictogram label condition, 12 male and 6 female students participated with the mean age of 21.3 years. All participants reported normal hearing and had normal or corrected to normal vision.

Stimulus materials

Auditory stimuli

Thirty-three product sounds were used. The sounds were taken from sound-effect CDs or were recorded using a recording apparatus (Boss BR-532, with a Sennheiser e865 microphone). Sounds longer than 5 seconds were trimmed to a maximum duration of 5 seconds (using Felt Tip Sound Studio v2.1). Sounds shorter than or equal to 5 seconds were not changed. The sounds were saved at CD quality and were presented at a similar comfortable listening level.

Of the 33 sounds, 23 were target sounds; 4 (two primacy and two recency sounds) were used to prevent primacy and receny effects (see Bartlett, 1977); and 6 were distracter sounds. Target sounds were presented throughout the whole experiment except the free recall task. They were chosen to represent the six product sound categories mentioned earlier (see Chapter 1). Five of the six product sound categories were represented by four target sounds; only the liquid sound category was represented by three target sounds. 'Primacy and recency' sounds ('PR' sounds) were also presented throughout the whole experiment. Distracter sounds—each representing one sound category—were presented only for the recognition task.

Labels

Each product sound had a corresponding label in any of the three formats (text, image, and pictogram). For each label type condition, a unique label was assigned to each sound. The distribution of labels and sounds was identical. Of the 33 labels, 23 were target labels; 6 were distracter labels; and 4 accompanied the 'PR' sounds. Target labels accompanied the target sounds in the label-sound presentation and matching task. Distracter labels were present only in the matching task.

*Text Labels****.*** The text labels were created using three guidelines: (a) the appliance produces only one type of sound, then the description of the appliance was followed by the operational state (e.g., vacuum cleaner: on); (b) the source of the sound is ambiguous, then the source and the action description were used (e.g., water boiling); (c) a part of an appliance causes the sound, then descriptions of this part, action, and appliance were used (e.g., microwave oven: door closing).

Image labels. Image labels were colour photos that were selected from product catalogues or made by the authors. The photos showed human-product interaction or only the device causing the sound. They were chosen using two guidelines: if the action was an important part in causing the sound, a photo was used that showed the hand of a person handling the product (e.g., rotating the button of a washing machine). In the other cases, the entire device was shown. The photos were digitized and sized to 283 x 283 pixels having a resolution of 72 dpi.

Pictogram labels. A set of pictograms was arranged for each sound and contained three to six pictograms depending on the main descriptive features of a sound. A pictogram set always contained a graphical depiction of a location where the sound could be heard, a graphical depiction of the amplitude of the sound as a function of time, and a graphical depiction of the source. Source descriptions were mostly the parts of the appliances causing the sound (e.g., fan, engine, etc.). For alarm sounds not the source but a graphical depiction of the meaning was used. For examples, see Figure 1.

Apparatus

The stimuli were presented using a specially designed application developed using the Trolltech Qt (Mac OS X - free edition) tool kit. The application ran on a Macintosh Powerbook G4 1.33 GHz computer with 12" screen having a resolution of 1024 x 768 pixels. The stimuli were presented through AKG Studio Monitor K240DF 2x600 Ohm headphones. The experiment took place in a quiet room.

Procedure

Each participant received a written explanation of the purpose of and the instructions for the study. A participant was seated in front of the screen at a distance of approximately 50 cm. The entire experiment was self-paced and there were no pauses between the different phases of the experiment. Two primacy sounds were always played at the beginning and two recency sounds at the end of each phase. The order of presentation of sound categories and of sounds within a sound category was randomly determined. Thus, sounds were always presented within their category.

In three of the four label type conditions (text, image, and pictogram), a label in the format of text, image, or pictogram was provided for each sound by the experimenter. In the self-generated text condition, a participant had to create a text label by typing it on the computer screen. Before the experiment started, participants in the pictogram condition were shown the pictograms and were explained how to interpret the sequence of pictograms. The experiment consisted of four phases: (1) sound-label presentation, (2) free recall, (3) recognition, and (4) matching for each label condition. Participants were informed of all the phases except the free recall phase.

In the *sound-label presentation* phase, the sound-label combination was presented to a participant. This was done in the following way: First, a participant had to listen to a sound, and second, a corresponding label was either presented on the screen in three of the four experimental conditions (text, image, and pictogram), or created by a participant in the self-generated text label condition. This procedure was repeated for each of the 27 sound-label combinations—23 target sounds plus the 4 ‘PR’ sounds. A participant was instructed to remember the sound-label combination presented or generated for the rest of the experimental tasks.

In the *free recall* phase, participants were asked to remember as many sounds as possible that were presented in the first phase. They were instructed to type a verbal text on the computer screen that described each sound they remembered. This phase was finished when a participant could no longer provide any descriptions.

In the *recognition* phase, 33 sounds—27 sounds plus 6 distracter sounds—were presented without any labels. The distracter sounds were included to test the ability to distinguish between previously presented sounds and new sounds. Participants were forewarned that distracter sounds were added in the list. A participant’s task was to listen to the presented sound and to rate it on a 6-point scale (‘1’ representing ‘I am sure this is a new sound’; ‘6’ representing ‘I am sure this is an old sound’). The randomization procedure was identical to the first phase. A participant could proceed to the *matching* phase after all the sounds were rated.

In the *matching* phase, all labels corresponding to each experimental condition were presented on a paper together with a two-digit code besides the keyboard. This provided a constant visibility throughout the whole phase. Participants in the self-generated text label condition were provided with the labels that they had created in the first phase (list presentation). Six additional labels that corresponded to the distracter sounds of the recognition phase were added as distracter labels in all conditions. Participants were forewarned about the distracter labels. In total 33 labels and 27 sounds were presented for the text, image, and pictogram label conditions. For the self-generated text label condition, all the sounds a participant managed to recall plus six distracter sounds were presented. A participant first listened to the sound and then tried to match it to the correct label. The code was typed in a text box on the screen. The randomization of the sounds was identical to the first phase. The experiment ended when all sounds were matched to a label.

	Label Type			
Tasks	**Self-Generated Text**	**Text**	**Image**	**Pictogram**
Free Recall	0.51 (0.07)	0.49 (0.08)	0.56 (0.07)	0.15 (0.08)
Recognition	5.00 (0.23)	4.87 (0.22)	4.69 (0.22)	4.38 (0.24)
Matching	0.57 (0.07)	0.48 (0.05)	0.56 (0.06)	0.34 (0.06)

Note. Numbers in parenthesis are standard errors of the mean.

Table 1. Mean Proportion Correct Responses for Free Recall and Matching Tasks and Mean Oldness Ratings for Recognition Task as a Function of Label Type

	Label Type			
Tasks	**Self-Generated Text**	**Text**	**Image**	**Pictogram**
Free Recall	0.43 (2)	0.31 (3)	0.74 (1)	-1.50 (4)
Recognition	0.99 (1)	0.50 (2)	-0.20 (3)	-1.30 (4)
Matching	0.78 (1)	0.00 (3)	0.66 (2)	-1.40 (4)
Mean Over Tasks	*2.20 (1)*	*0.76 (3)*	*1.24 (2)*	*-4.20 (4)*

Note. Numbers in parenthesis represent the ranking of label type for each task.

Table 2. Z-Scores of Mean Proportion Correct Responses for Free Recall and Matching Tasks and Mean Oldness Ratings for Recognition Task as a Function of Label Type.

	Sound Type					
Tasks	**Air**	**Alarm**	**Cyclic**	**Impact**	**Liquid**	**Mechanical**
Free Recall	1.04 (1)	0.55 (3)	0.31 (4)	-1.52 (6)	-0.95 (5)	0.55 (2)
Recognition	-0.89 (6)	1.70 (1)	0.11 (3)	-0.77 (5)	-0.66 (4)	0.52 (2)
Matching	-0.57 (5)	1.52 (1)	-1.40 (6)	-0.31 (4)	0.43 (2)	0.32 (3)
Mean Over Tasks	*-0.42 (3)*	*3.78 (1)*	*-.0.98 (4)*	*-2.60 (6)*	*-1.18 (5)*	*1.39 (2)*

Note. Numbers in parenthesis represent the ranking of sound type for each task.

Table 3. Z-Scores of Mean Proportion Correct Responses for Free Recall and Matching Tasks and Mean Oldness Ratings for Recognition Task as a Function of Sound Type

Results

The four sounds used to prevent the primacy and recency effect were excluded from the analysis. For the free recall, recognition, and matching phases, the data were analyzed with an ANOVA with label type as the between-subjects factor (4 levels) and sound type as the within-subject factors (6 levels).

In *free recall,* three types of responses were observed: correct responses that semantically or syntactically matched the target sounds; incorrect responses which semantically mismatched the target sounds; and no response. The sounds with correct responses were scored as '1', with incorrect responses as '-1', and not recalled sounds as '0'. The sums were divided by the number of sounds within a sound category. Table 1 presents the proportion correct for free recall as a function of label type. In the table, image condition has the highest proportion correct (.56), and the pictogram condition the lowest (.15). Figure 2 presents the proportion correct for free-recalled sounds as a function of label type and sound type. In the figure, the impact and liquid sounds have the lowest proportion correct (.30) over all conditions; the other sound categories score higher (~0.50). Significant effects for label and sound types were found, $F(3, 68) = 24.58$, $p<.001$, and $F(5, 340) = 5.91$, $p<.001$, respectively. An interaction effect was found for the label types and sound categories, $F(15, 340) = 1.73$, $p<.05$. Air and liquid sounds were better recalled in the self-generated text condition that in the provided text condition. In addition, alarm and cyclic sounds are better recalled in the provided text condition that in self-generated text condition. A post-hoc analysis was conducted to reveal which levels of the main effects differed significantly. For label type, only pictograms differed significantly from the other label types ($p<.001$). For sound type, impact sound differed significantly from air and alarm sounds ($p<.001$) and differed significantly from cyclic and mechanical sounds ($p<.01$); liquid sounds differed significantly from air, alarm, and mechanical sounds ($p<.05$).

The free recall performance as a function of label type (Table 1) decreased in the following way: image, self-generated text, text, and pictogram. The free recall performance as a function of sound type (Figure 2) decreased in the following way: air, alarm, mechanical, cyclic, liquid, and impact.

Recognition score was determined for target and distracter sounds. The means of the ratings for target and distracter sounds confirmed that target sounds were rated as old (the mean of the ratings ranged between 3.7 and 5.4) and distracter sounds as new (the mean of the ratings ranged between 1.7 and 4.2). In Table 1, mean

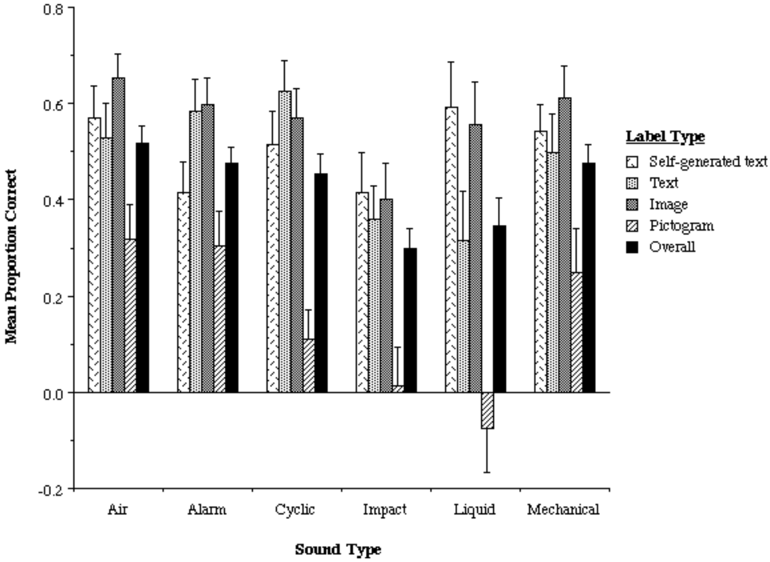


Figure 2. The mean proportion correct responses for free recall task as a function of sound type and label type is shown. The error bars in the y-axis represent the standard error of the mean.

recognition rating (oldness) of a sound was highest for self-generated text label condition and lowest for pictogram label condition. Figure 3 shows that the mean oldness rating was highest for alarm sounds (5.2) and lowest for air, impact, and liquid sounds (~4.6). Significant effects for the label and sound types were found, *F*(3,68) = 3.67, *p*<.05 and *F*(5, 340) = 5.55, p<.001, respectively. In addition, no interaction effect was found between the label types and sound categories, *F*(15, 340) = 1.67, NS. A post-hoc analysis was conducted to reveal which levels of the main effects differed significantly. For label type only pictograms differed significantly from self-generated text and text (*p*<.05). For sound type, alarm sounds differed significantly from air sounds, impact, and liquid sounds (*p*<.001), and from cyclic and mechanical sounds (*p*<.05); and mechanical sounds differed significantly from air, alarm, impact and liquid sounds (*p*<.05).

The recognition performance as a function of label type (Table 1) decreased in the following way: self-generated text, text, image, and pictogram. The recognition performance as a function of sound type (Figure 3) decreased in the following way: alarm, mechanical, cyclic, liquid, air, and impact.

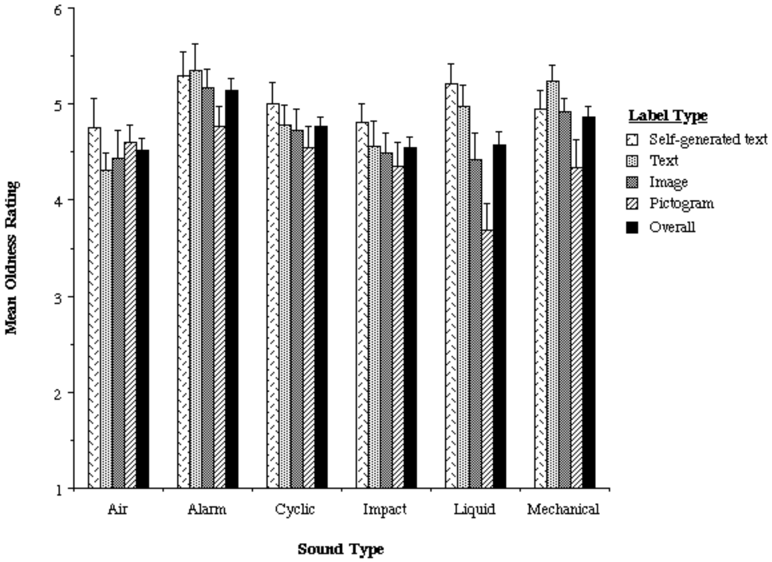


Figure 3. The mean oldness rating for recognition task as a function of sound type and label type is shown. The error bars in the y-axis represent the standard error of the mean.

Matching responses were scored as '1' if a correct label was assigned to a sound, otherwise, as '0'. Table 1 presents the proportion correct of the matchings for the sounds as a function of label type. In the table, proportion correct is lowest for the pictogram condition (.34), and highest for the self-generated text condition (.57). In Figure 4, proportion correct varies across sound type; it is highest for the alarm sounds (.74) and lowest for cyclic sounds (.26). Significant effects for the label and sound types were found, $F(3,68) = 8.85$, $p<.001$ and $F(5, 340) = 38.12$, $p<.001$, respectively. In addition, no interaction effect was found between the label types and sound categories, $F(15, 340) = 1.42$, NS. A post-hoc analysis was conducted to reveal which levels of the main effects differed significantly. For label type only pictograms differed significantly from self-generated text and image ($p<.001$), and differed significantly from text ($p<.05$). For sound type, alarm sounds and cyclic sounds differed significantly from the other sound types ($p<.001$); and impact sounds differed significantly from alarm and cyclic sounds ($p<.001$) and from liquid and mechanical sounds ($p<.05$); and liquid sounds differed significantly from air, alarm, and cyclic sounds ($p<.001$) and from impact sounds ($p<.05$).

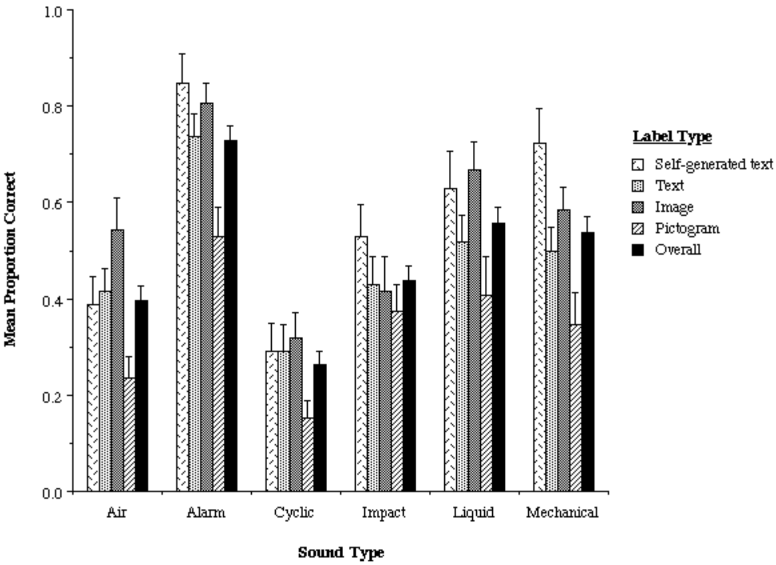


Figure 4. The mean proportion correct responses for matching task as a function of sound type and label type is shown. The error bars in the y-axis represent the standard error of the mean.

The matching performance as a function of label type (Table 1) decreased in the following way: self-generated text, image, text, and pictogram. The matching performance as a function of sound type (Figure 4) decreased in the following way: alarm, liquid, mechanical, impact, air, and cyclic.

In order to make a more general memory overview, a combined measure to reflect general memory performance over all tasks (free recall, recognition, and matching) was developed. The proportion correct scores for free recall and matching tasks and recognition ratings were transformed into Z-scores**. The mean of these scores was calculated for label types and sound types separately. Table 2 presents the Z-score means for label type. The means determined the ranking order (indicated in parenthesis). This ranking order shows that memory performance decreased in the following way for label type: self-generated text, image, text, pictogram. Table 3 presents the Z-score means for sound type determining the ranking order. For sound type, the following order was observed for decreasing memory performance: alarm, mechanical, air, cyclic, liquid, and impact sounds.

Discussion

Sound type as well as label type affect the memory performance for product sounds. Sound type affects memory in the predicted way. That is, the amount of structure in the spectral and temporal domains is the determinant factor. Structured and moderately structured sounds (alarm and mechanical sounds) are better remembered in the three memory tasks. Unstructured sounds (liquid and impact sounds) mostly have worse memory performances for free recall and recognition tasks, whereas no such effect has been found for the matching task. The type of label presented together with a sound also affects the memory performance depending on the task and the label type. We have observed two types of overshadowing effects for the recognition task. One is the verbal overshadowing effect seen in the provided text labeling as opposed to the self-generated text labeling. The other is the visual overshadowing effect resulting from the density and complexity of the provided visual information at encoding. Whereas pictogram labeling suffers the most from visual overshadowing effect in all tasks, self-generated text labeling enhances the memory performances especially in recognition and matching tasks. Furthermore, sounds presented with image labels have been better remembered in free recall and matching tasks due to dual-coding. In the following sections these findings will be discussed more elaborately.

The effect of sound type

The amount of structure in a product sound determines how easily a sound can be encoded and reproduced. This is especially evident for the free recall and recognition tasks. These memory tasks do not make use of additional perceptual cues during retrieval and comparison phases in a memory task, but do mostly rely on the sound's structural properties. Free recall highly depends on the ability to reproduce the target sounds through an internal search. Our results show that free recall performance is higher for sounds that consist of a consistent structure. This may indicate that the ability to reproduce the sound in mind requires accessibility to the structure of the previously coded sound. Once the reproduction of the sound is complete, it becomes easier to activate the sound's semantic associations. Therefore, accessibility to the structure of the sound and the ability to reproduce the structure may explain the better memory performance for structured sounds. Similarly to Bartlett's study (1977), a recognition task can be considered as a process in which a previously derived structure of a sound in memory is compared to the derived structure of the sound just heard. The fit between these structures will determine the level of recognition. The predicted ranking of product sounds based on the spectral-temporal structure shows that the structural aspects are important in recognition memory. Only the order of liquid and air sounds has been changed.

Alarm sounds, which consist of a tonal spectrum and a rhythmic pattern, have been better remembered than mechanical sounds, which consist of some harmonic bands and noise in the spectrum and have a temporal periodicity. Although cyclic sounds resemble air sounds in terms of spectral composition (noise-like sounds with harmonic bands), they have a more specific temporal periodicity that makes them more distinguishable than air sounds. Liquid and impact sounds, being noise-like and having no temporal regularity, are to be remembered with least accuracy, because listeners will fail to derive the spectral-temporal structures. However, changing spectral structure of liquid sounds over time may be indicative of an event phenomenon and slightly improve the memory for such sounds as opposed to impact and air sounds.

The superior effect of the structural composition is less for the matching task. Structurally irregular sounds (liquid and impact sounds) have been better matched than moderately structured (mechanical, cyclic and air sounds) and even well-structured sounds (alarm sounds). The reason for this may be that a matching task is procedurally different compared to the free recall and recognition tasks. In the employed matching task, both the target sound and the target label are available. Consequently, additional perceptual and semantic cues have been provided which probably moderated the cognitive processes for the match. In this type of matching task, structural analysis of the sound may not be the only essential factor. The structural analysis may be followed by accessing the conceptual information on the sound and matching it to the conceptual information that any of the labels evoke. Thus, this process may be beneficial for sounds with irregular structure because the incomplete encoding caused by irregular structure may be compensated by the presence of the additional semantic information. Furthermore, this may imply that the superiority effect of the structure is more evident in the cases in which additional perceptual and semantic information for the to-be-remembered target sound is absent.

The effects of label type

Self

The results have shown a consistent advantage for self-generated labels and a significant disadvantage for pictogram labels. It seems plausible that memory performance is higher when participants generate their own labels rather than when the labels are provided. Greenwald and Banaji (1989) showed that recall accuracy of target nouns increased for sentences that included familiar names (i.e., names of friends). This mnemonic benefit has been explained by “self’s being a highly familiar and rich knowledge structure”. Thus, previously experienced events/objects are assumed to create strong cues and associations that are beneficial for cognitive functions of related objects. This indicates that self is able to easily associate knowledge related to the own past experiences and to the self. Similarly, when participants label sounds, the meaning derived from an auditory event is based on their own words, or idiosyncratic labels that are used. Thus, the active involvement of self in the encoding process seems to have facilitated the retrieval process for free recall, recognition, and matching tasks.

For memory tasks that incorporate the retrieval of labels (i.e., free recall and matching), it is plausible that self-generated labels would have superiority. However, interestingly the same beneficial effect has been observed for the recognition task. One suggestion is that if identification occured during the recognition task, it would have been easier for the participants in self-generated label condition to identify the sounds due to self's bias (Cleary, 2002; Greenwald & Banaji, 1989).

Verbal and visual overshadowing

The negative effects of labeling of complex stimuli for recognition are previously discussed in other studies in terms of verbal and perceptual discrimination of voices or wine tastes (Schooler & Engtsler-Schooler, 1990; Melcher & Schooler, 1996; Perfect et al., 2002). These studies mainly focus on the verbal label as the interfering factor and altered the type of complex stimuli (i.e., auditory, olfactory, visual). In this study, two types of overshadowing effects have been observed. One is the verbal overshadowing effect caused with the text labels, which is similar to the previous studies. The other is the 'visual' overshadowing effect, presumably caused by the density of the visual information presented as labels for product sounds.

The overall recognition ratings for each label condition reveal that recognition accuracy suffers from the density of information. At encoding, participants were provided with the least amount of information in the self-generated text condition—they had to provide a label themselves. However, in the other conditions (*a*) the semantic association of the label was directly presented to the participant (i.e., text label), or (*b*) participants had to first access the semantic associations of the given label and then relate the relevant association to the sound (i.e., image and pictogram conditions). Participants who generate the text labels themselves are able to focus on encoding information in the auditory modality, whereas participants in the provided text label condition have to focus on auditory modality plus the given semantic association. This may have caused the 'verbal' overshadowing effect.

In addition, encoding auditory information with the presence of visual labels may be a challenging task for a listener, because the semantic association of the label is not apparent in its relation to the sound. Furthermore, the relation between the auditory and the visual information has to be established well. Thus, the disadvantage of the visual labels (i.e., image and pictogram) is that the participants' attention was split over two modalities (auditory and visual) in the encoding phase. This may have caused the ‘visual’ overshadowing effect. Therefore, in general, the amount and type of information presented as labels may have substantially interfered in encoding the auditory content of the sound (e.g., spectral-temporal structure).

Free recall and matching accuracies suffer significantly from the visual overshadowing effect caused by the pictogram labels. Probably ambiguous and complex sounds have an additional disadvantage to evoke more than one semantic association. As a result, lower scores in the recall phase are expected because there is no concrete label to recall. Initially at encoding, the number of associations may have inhibited to assign one semantic association. This probably results from the activation of too many nodes at encoding due to the variety in a pictogram set—a pictogram set contained at least four (maximum six) different types of pictograms.

Dual coding

Although the complexity of the labels has a negative effect especially on the recognition performance, free recall and matching performances benefited from dual coding of the sounds together with images. According to the dual coding theory, images are mnemonically superior to verbal codes because they are high in imagery. Moreover, confirming Thompson and Paivio’s (1994) study, encoding of dual-modality information (auditory and visual codes) has been advantageous especially for the free recall task. Because of the serial encoding of the auditory and visual information, the conceptual information that corresponds both to the sound and the image must have been activated twice. As a result, the bonds between the stores of the two perceptual systems get stronger. Such a double encoding must have enabled the easy access to the perceptual or semantic stores at the retrieval phase during free recall. Participants who are able to reproduce a sound probably accessed its label (semantic association) via the auditory store. However, double encoding may have been advantageous in the cases that participants failed to reproduce the sound (due to the ambiguity or complexity) and to provide a consequent label. Then, visual codes may have helped to access to the conceptual information that corresponds to the sounds and provide a label.

A more general memory measure?

The prediction of the general memory performance is supported by the combined measurement reflecting general memory performance over all tasks. As mentioned before, the overall ranking of the memory performance for sound type is as predicted. Only air and cyclic sounds have inter-changed rankings, the reason for which is difficult to explain. However, the inter-change may be the result of the observed similarity between air and cyclic sounds in a categorization task (see Chapter 1). This may imply that the perceptual similarity of the sounds may have interfered the ranking. Moreover, only the position of the two sound groups has been changed. Consequently, the structural properties of product sounds are predictive of the memory for product sounds. Considering the dependence on the label type conditions, the general ranking is similar to the prediction. However, the ranking fails to distinguish between the verbal and visual labels. Self-generated text labels score better than the provided text labels and image labels score better than the pictogram labels, whereas self-generated text labels have the superiority to all labels and image labels score better than the text labels. This implies that the effect of different type of labels on product sound memory is independent of modality effects.

General conclusions

This study has shown that the memory of product sounds—which are ambiguous and complex in nature—can be predicted by the spectral-temporal structure present in a sound. Another main finding of this study is that the memory performance for label type is inconsistent over tasks. This may indicate that memory performance for product sounds is task-dependent. Each task requires a different type of processing. For example, a recognition task requires structural analysis that accesses the codes in the perceptual store in order to reproduce the item; or free recall task requires the retrieval of the label from the semantic store. Depending on the requirements of the task, the memory system makes use of the necessary information to maximize the performance.

The visual overshadowing effect has been proven robust in all memory tasks. This robust effect may be a result from the high perceptual expertise on the auditory domain (product sounds are somewhat familiar environmental sounds) and insufficient perceptual expertise on the visual domain (pictographic language as a concept is relatively new to the participants). Previous studies (Melcher & Schooler, 2004; Tindall-Ford et al., 1997) have suggested that with conceptual training a verbal overshadowing effect can be diminished. Moreover, consistent with the literature, the results suggest that when the complexity of the visual information is reduced, or when the expertise is high in both auditory and verbal domain, or auditory and visual domain (i.e., pictures in this case), the recall performance increases as a result of dual-coding theory. Therefore, the future focus of this study will be on the learnability of the pictograms and whether with perceptual expertise on the visual labeling the visual overshadowing effect will be minimized. For this, (sound) design practitioners should be selected as a sample group. Another attempt will also be made on the simplification of the pictographic description to decrease the amount of visual information at encoding.

Furthermore, the negative effect of complex visual information on the encoding and recognition of the information from other modalities should be tested further in contextual situations in which visual objects may create an overload of visual information. To start with, the memory for product sounds can be tested in living environments to which they belong to investigate whether visual objects in that context will create the same visual overshadowing effect.

Acknowledgements

We would like to thank Ilona de Jongh, Guido Hermans, Floris Schrader for running of the experiments, Ans Koenderink-van Doorn for the useful discussions, and Alexandre Vieira for voluntarily programming the experimental software in his free time. We gratefully acknowledge the valuable comments of Joseph Melcher during the review process of the paper.

Footnotes

* In some fields of research, the term ‘device’ is used instead of ‘product’. In the field of industrial design engineering, the term ‘product’ is mostly commonly used.

** Z-scores were employed because of free recall, matching, and recognition tasks were measured on different scales.

References

Bahrick, H. P., & Boucher, B. (1968). Retention of visual and verbal codes of the same stimuli. *Journal of Experimental Psychology 78*(3), 417-22.

Bartlett, J. C. (1977). Remembering environmental sounds: The role of verbalization at input. *Memory & Cognition 5*(4), 404-414.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding*. Psychological Review 94*(2), 115-47.

Carmichael, L., Hogan, H.P., & Walters, A.A. (1932). An experimental study of the effect of language on the reproduction of visually perceived form. *Journal of Experimental Psychology. 15*, 73-86.

Chiu, C. Y. P. & Schacter, D. L. (1995). Auditory priming for nonverbal information: Implicit and explicit memory for environmental sounds. *Consciousness and Cognition, 4*(4), 440-458.

Cleary, A. M. (2002). Recognition with and without identification: Dissociative effects of meaningful encoding. *Memory & Cognition, 30*(5), 758-767.

Deutsch, D. (1972). Mapping of interactions in the pitch memory store. *Science, 175*(25), 1020-1022.

Deutsch, D. (1980). The processing of structured and unstructured tonal sequences. *Perception & Psychophysics, 28*(5), 381-9.

Deutsch, D., & Feroe, J. (1981). The internal representation of pitch sequences in tonal music*. Psychological Review, 88*(6), 503-522.

Edworthy, J., & Hards, R. (1999) "Learning Auditory Warnings: The Effects of Sound Type, Verbal Labeling and Imagery on the Identification of Alarm Sounds." International Journal of Industrial Ergonomics 24: 603-618.

Garner, W. R. (1974). The perception and learning of temporal patterns. In *The processing of information and structure* (pp. 49-72). New York: Wiley.

Greene, A. J., Easton, R. D., & LaShell, L. S. R. (2001). Visual-auditory events: Cross-modal perceptual priming and recognition memory. *Consciousness and Cognition, 10*(3), 425-435.

Greenwald, A. G., & Banaji, M. R. (1989). The self as a memory system: Powerful, but ordinary. *Journal of Personality and Social Psychology, 57*(1), 41-54.

Huss, M. T., & Weaver, K. A. (1996). Effect of modality in earwitness identification: Memory for verbal and nonverbal auditory stimuli presented in two contexts. *Journal of General Psychology, 123*(4). 277-287.

Lehrer, A. (1983). *Wine and conversation*. Bloomington: Indiana University Press.

Liu, T. S., & Cooper, L. A. (2001). The influence of task requirements on priming in object decision and matching. *Memory & Cognition, 29*(6), 874-882.

Lyman, B. J., & McDaniel, M. A. (1990). Memory for odors and odor names: Modalities of elaboration and imagery. *Journal of Experimental Psychology: Learning, Memory and Cognition, 16*(4), 656-664.

Martens, W., & Giragama, C. (2002). *Relating multilingual semantic scales to a common timbre space*. Proceedings of 113th Convention of the Audio Engineering Society.

Melcher, J. M., & Schooler, J. W. (1996). The misremembrance of wines past: Verbal and perceptual expertise differentially mediate verbal overshadowing of taste memory. *Journal of Memory and Language, 35*(2), 231-245.

Melcher, J. M., & Schooler, J. W. (2004). Perceptual and conceptual training mediate the verbal overshadowing effect in an unfamiliar domain. *Memory & Cognition, 32*(4), 618-631.

Özcan, E., & van Egmond, R. (2004). Pictograms for sound design: A language for the communication of product sounds. *Proceedings of the International Conference on Design and Emotion, Turkey, 4,* 1-16.

Özcan, E., & van Egmond, R. (2006) Product sound design and application: An overview. *Proceedings of the International Conference on Design and Emotion*, *Sweden, 5,* 1-19.

Özcan, E., van Egmond, R., & Jacobs, J. (submitted). Bases for categorization and identification of product sounds.

Paivio, A. (1983). The empirical case for dual coding. In J. C. Yuille *(Ed.), Imagery, memory and cognition: Essays in honor of Allan Paivio*. Hillsdale: Lawrence Erlbaum Associates.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology - Revue Canadienne De Psychologie, 45*(3), 255-287.

Paivio, A., Philipchalk, R., & Rowe, E. J. (1975). Free and serial recall of pictures, sounds, and words. *Memory & Cognition, 3*(6), 586-590.

Palmer, S. E. (1977). Hierarchical structure in perceptual representation. *Cognitive Psychology, 9*(4), 441-474.

Perfect, T. J., Hunt, L. J., & Harris, C. M. (2002). Verbal overshadowing in voice recognition. *Applied Cognitive Psychology, 16*, 973-980.

Povel, D. J. (1981). Internal representation of simple temporal patterns. *Journal of Experimental Psychology: Human Perception and Performance, 7*(1), 3-18.

Schacter, D. L., Cooper, L. A., & Delaney, S. M. (1990). Implicit memory for unfamiliar objects depends on access to structural descriptions. *Journal of Experimental Psychology: General, 119*(1), 5-24.

Schooler, J. W., & Engstler-Schooler, T. Y. (1990). Verbal overshadowing of visual memories: Some things are better left unsaid. *Cognitive Psychology, 22*(1), 36-71.

Sternberg, R. J. (1998). *Cognitive psychology* (2nd Ed). Fort Worth: Harcourt Brace.

Sweller, J., & Chandler, P. (1991). Evidence for cognitive load theory. *Cognition and Instruction, 8*(4), 351-

362.

Sweller, J., & Chandler, P. (1994). Why some material is difficult to learn. *Cognition and Instruction, 12*(3), 185-233.

Thompson, V. A., & Paivio, A. (1994). Memory for pictures and sounds: Independence of auditory and visual codes. *Canadian Journal of Experimental Psychology, 48*(3), 380-398.

Tindall-Ford, S., Chandler, P., & Sweller, J. (1997). When two sensory modes are better than one. *Journal of Experimental Psychology: Applied, 3*(4): 257-287.

Vanags, T., Carroll, M., & Perfect, T. J. (2005). Verbal overshadowing: A sound theory in voice recognition? *Applied Cognitive Psychology, 19*: 1127-114.

Van Egmond, R. (2006). Designing an emotional experience for product sounds. *Proceedings of the International Conference on Design and Emotion, Sweden, 5*, 1-12.

This chapter is based on the paper:

Özcan, E., & van Egmond, R. (2007). ***How well do we identify product sounds? Proceedings of the International Conference on Auditory Display, Canada. 13,*** **234-241**

Abstract

The lexical associations for a set of 29 product sounds were determined in two experiments. Experiment 1 showed that listeners fail in correctly identifying a product sound in a free identification task and naming errors occur during labeling because of high perceptual similarities. Experiment 2 investigated the number and variety of lexical associations a product sound may have in semantic memory and determined the causal uncertainty values for product sounds. The results indicate that product sounds are lexically not well represented in memory and that identification accuracy decreases with high causal uncertainty. Findings suggest that auditory information from product sounds may be semantically represented in memory, but for some sounds these representations are fuzzy and not easily accessible.

CHAPTER 4

⁂

HOW WELL DO WE IDENTIFY PRODUCT SOUNDS?

The sound of a product can be informative about the parts and actions involved in the functioning of the product. Correct auditory identification influences how the product is experienced. For example, misidentifying the sound of a toothbrush as a dentist drill will retrieve memories about a dentist context and therefore may cause an unpleasant experience. Identifying a rotating brush sound of the toothbrush as a moving blade of a shaver will also influence the consequent actions that a user may take. Thus, attribution of meaning caused by auditory information is important in product-user interaction. However, correct auditory identification may be a difficult task for users (i.e. listeners) as products emit perceptually similar sounds (e.g., an electric toothbrush, shaver, and hair clippers produce acoustically similar sounds). Therefore, the extent to which a product sound is identified needs to be investigated. Thus, the cognitive and acoustical factors that may take place during an environmental sound identification will first be discussed.

Ambiguity and causal uncertainty

Most of the studies regarding environmental sound identification have so far focused on the processing of auditory information on a perceptual or a cognitive level (Ballas, 1993; Bonebright, 2001; Bregman, 1990; Fabiani, Kazmerski., Cycowicz, & Friedman, 1996; Gygi, Kidd, & Watson, 2004; Handel, 1991; Marcell, Borella, Greene, Kerr, & Rogers, 2000; McAdams & Bigand, 1993). These studies have investigated listeners' ability to identify, label, and to categorize environmental sounds and have provided insight into the sounds' semantic associations in memory (see also, Özcan & van Egmond, 2005; Vanderveer, 1979). Some studies directly measured the identifiability degree of the environmental sounds and the response time needed to label the cause of the sound (Ballas, 1993; Ballas & Mullins, 1991; Guillaume, Pellieux, Chastres, Blancard, $ Drake, 2004; Vanderveer, 1979). However, the cause why certain sounds are more identifiable than others is not well known. Ballas' studies (Ballas, 1993; Ballas & Mullins, 1991) have shown that causal uncertainty, namely, ambiguity may cause difficulties in sound identification.

Ambiguity in sound identification may occur if a sound has multiple causes. For example, an old-fashioned alarm clock, a kitchen timer, a clockwork toy, or a school bell may cause the same high-pitched, continuous, rattling sound. Although the sounds are perceptually very similar, the causes of the sounds are contextually dissimilar. Thus, auditory information from such sounds may be represented individually in semantic memory and have different lexical associations (Özcan & van Egmond, 2005). This may create confusions in accurate sound labeling, because an ambiguous sound can potentially activate more than one lexical association in memory. Accordingly, memory representations play an important role in correct identification.

Memory representations

Auditory memory is capable of storing auditory information per se (Bartlett, 1977, Deutsch, 1980; Crowder, 1993) and it is also linked to other perceptual or semantic stores via conceptual associations (Paivio, 1991; Thompson & Paivio, 1994; Özcan & van Egmond, 2007). General findings are that memory favours hierarchical units in the structure of a sound and auditory information is able to activate a label, but not vice-versa. Özcan & van Egmond (2007) have investigated the recognition, free recall and matching memory for product sounds. It has been shown that spectral-temporal structure in a product sound can be predictive of good memory performance and the memory performance for product sounds is task-dependant. For a recognition task, which requires perceptual analysis and comparison, encoding product sounds without text or image labels seems to be the most beneficial. Consequently, because of verbal and visual overshadowing effects, recognition performance decreases as the semantic information at encoding increases from no-label to text label and image labels. For free-recall and matching task, which require conscious recall of the name of the sound, encoding sounds with image labels is the most beneficial as a result of the dual coding.

Labeling

A commonly used method for measuring sound identification accuracy is free labeling (Ballas, 1993; Berglund & Nilson, 2003; Fabiani et al., 1996; Marcell et al., 2000; Vanderveer, 1979). Such a paradigm allows listeners to describe a sound without any constraints. Studies, which used this paradigm, have shown that listeners primarily tend to describe the cause (i.e., source and action descriptions) of the sound rather than the acoustical properties. Acoustical and structural properties are described when no identification occurs (Handel, 1991; Vanderveer, 1979). Identification accuracy is operationalized as correct when it semantically matches the label of the cause (e.g., door closing) (Marcell et al., 2000).

A free labeling paradigm produces other semantic associations that a sound may possibly have—apart from the cause of the sound. Fabiani et al. (1996) have categorized such descriptions as not-known (e.g., disgusting noise), sound imitation (e.g., too-too-too), sound description (e.g., high-pitched), name or compound name (e.g., bird, water drain bubbles). They also determined the level of the conceptual association (car for modal, automobile for synonym, truck for coordinate, vehicle for super-ordinate, Ferrari for subordinate). Özcan & van Egmond (2005) have indicated that product sounds are represented on 11 different levels of semantic associations (i.e., source, action, onomatopoeias, emotion, source properties, psychoacoustics, material, location, temporal aspects, abstract meanings, and emotional responses).

Perceptual similarity and categorization

Although sound descriptions provide an extensive insight into the semantic, or more precisely, verbal associations of sounds, they still cannot categorically distinguish similar sounds. However, perceptual similarity may play an important role in assigning the correct name to a sound. Environmental sounds may have (a) structural similarity when they share similar spectral-temporal composition but are semantically dissimilar, such as old-fashioned alarm clock and a kitchen timer, (b) semantic similarity when they share a similar name but are structurally dissimilar, such as an old-fashioned and a digital alarm clock, and (c) contextual similarity when they co-occur in natural scenes, such as kitchen timer and kitchen hood sounds, or washing machine sound and a washing machine rotary button sound. Therefore, studies have investigated on what ground listeners find similarities between sounds and categorize them (see, Handel, 1991). Gaver, excluding musical or speech sounds (1993) has proposed that interacting objects can be theoretically discerned into three main classes of sound producing events (i.e., vibrating objects, aerodynamic sounds, liquid sounds) based on the material structure of the object, type of action, and the medium in which they are produced.

Special methods such as perceptual-cognitive rating or free categorization have been employed to define similarities between environmental sounds. For example, Ballas (1993) has concluded that listeners' similarity judgments are based on the perceptual dimensions (e.g., timbre), which also reflect particular type of events. Marcell et al. (2000) used a free categorization paradigm in which a category was assigned to sound while the sound was being identified. They have concluded 27 categories varying on the basic, sub- and super-ordinate concepts such as: locations (bathroom, kitchen), events (accident, sleep), objects (weapon, paper), creatures (animal, bird), situations (sickness), etc. The study of Özcan & van Egmond (2005) allowed participants compare the sounds with each other and label each category they created in a free categorization study. This study resulted in six product sound groups that vary in their spectral-temporal structure across categories: air, alarm, cyclic, impact, liquid, and mechanical sounds. The category labels revealed that similarities were based on (a) perceptual similarity (e.g., psychoacoustics, onomatopoeias, temporal descriptions), (b) cognitive similarity (e.g., sound source, location, abstract meanings), and/or (c) affective similarity (basic emotions). The studies above have shown that categorization may occur on different levels of concepts, thus, there may be fuzzy boundaries between categories. Moreover, perceptual judgments on the spectral-temporal structure of the sounds still guide the categorization process.

Meaning and spectral-temporal structure

Frequency content of a sound and how it changes over time can be informative about the object and the event causing the sound. Studies have shown that listeners can hear the material (Hermes, 1998; Klatzky, Pai, & Krotkov; 2000), shape (Kunkler-Peck & Turvery, 2000; Lutfi, 2001) of the object and the event (Aljishi, 1991; Cabe & Pittinger, 2000; Li, Logan, & Pastore, 1991) causing the sounds. Other studies have shown that changes in the timbre or rhythmic pattern of abstract sounds influence listeners' perceptual (sharpness, roughness) and emotional (obtrusive, unpleasant) judgments, or their judgments in more abstract concepts (urgency, danger) (Björk, 1985; Edworthy, Hellier, & Hards, 1995; Kandall & Carterette, 1993; Solomon, 1958; von Bismarck, 1974). Similarly, Gygi et al. (2004) have demonstrated that listeners are very sensitive to the auditory information and slight changes in the spectral-temporal content of the sound may influence the outcome of the identification process (i.e. labeling). Ballas (1993) has indicated that identification of sound is not only dependent on the spectral-temporal structure but also familiarity, ecological frequency, and other conceptual associations. Coward and Stevens (2004) have shown that the same sound with a concrete association (nomic mapping) is better recognized than same sound with an abstract association (symbolic mapping). The studies above suggest that bottom-up processing (i.e., perceptual analysis) is important for extracting meaning from sounds and the cognitive system makes use of the most plausible association.

Auditory identification process

Auditory identification is a complex process which incorporates a variety of perceptual and cognitive functions that any sound has to undergo (Bregman, 1990; Handel, 1991; McAdams & Bigand, 1993). For the identification to occur a sound has to pass through a recognition phase following the perceptual analysis phase (McAdams & Bigand, 1993). Recognition occurs if the results of the perceptual analysis of a sound match with any previously stored auditory codes (namely, mental representations). This phase is very crucial for building conceptual associations in memory, as identification should be completed by accessing to at least a semantic association and possibly to a lexical association. Cummings et al. (2006) have indicated that accessing to the meaningful semantic representation occurs before accessing to lexical representations. However, if no recognition occurs, then listeners can only describe the results of the perceptual analysis, namely, the spectral-temporal structure of the sound (Handel, 1991; Özcan & van Egmond, 2005; Vanderveer, 1979)

Studies, which measured identification accuracy for environmental sounds, have shown that listeners can accurately identify environmental sounds; this process favours rhythmic sounds which are as short as 150 ms (Guillaume, et al. 2004). Similarly, Vanderveer (1979) has shown that temporal pattern and high-frequency are determinants of perceptual identification and confusion occurs for impact sounds and for temporally similar sound. Ballas (1993) has shown that the processing time for the perceptual or cognitive analysis varies for different type of sounds.

Summary

There may be two explanations for causal uncertainty in environmental sound identification both stemming from high perceptual similarity between the sounds and both dependent on the recognition phase. First, perceptual analysis process may not always result in recognition. However, listeners have the tendency to attribute meaning to sounds. Thus, using the spectral-temporal structure, listeners may try to map this information to other perceptually similar sounds. This mapping may then yield several lexical associations. Secondly, perceptual analysis may result in recognition indicating that the sound is already represented in memory (access to semantic associations). It is possible that a single sound is represented with various concepts and has different lexical associations, which makes the cause of the sound ambiguous. In such situations, where ambiguity occurs, contextual cues may guide the identification process by limiting the number of possible causes. However, in the absence of context causal uncertainty occurs, because there are too many possibilities to choose from.

Similarly, this study will investigate the identifiability degree of a specific type of environmental sounds, namely product sounds. Products intrinsically produce similar sounds because they are built with standard parts (e.g., engines, fans, gears etc.) which are part of certain actions (e.g., rotating, sucking, impacting, etc.). Thus, we suspect that such sounds are low identifiable because of high causal uncertainty. We aim to provide insight into how well product sounds are lexically represented.

Experiments

Although the literature so far seems to be sufficient to derive conclusions for the identification process for product sounds, the domain of environmental sounds would still be too large focus to adopt the relevant information to product sound domain. The reasons are the following:

First, the environmental sound domain incorporates various domains of sounds such as speech or musical sounds, sounds caused by animals or natural events such as wind or rain, synthesized sounds, etc. The product sound domain is, however, one of the sub-domains. The domain comprises specific type of environmental sounds that result from the functionality of domestic appliances. Some examples are the sound of the hairdryer, dishwasher, shaver, coffee maker, toaster, and microwave oven finish beep.

Secondly, as the field of product design is developing, designers have started to put more focus on the sound design of the product (Özcan & van Egmond, 2006; van Egmond, 2006). This new trend requires new tools and methods to support the communication of the design team on this very specific field. For that, we (Özcan & van Egmond, 2006) have started to develop a special software by which designers can auditorily model their ideas—analogical to the 3D modeling programs—and present them to the design team. The sounding output of this software can eventually be used for the sound quality evaluation. Sound quality evaluation as a method employs semantic differential technique to assess the semantic associations that a sound may represent. As the listeners should focus only on the auditory information for better assessment, this method traditionally includes only the sound of a product for assessment, not the visual representation of it. Then, the activated semantic association depends solely on the auditory information. For this, we need to know whether product sounds are identifiable per se in the absence of visual information.

Moreover, auditory displays often employ alarm sounds, impact sounds, or sounds that refer to real events that may involve products (Gaver, 1989; Gaver, 1993; Keller & Stevens, 2004). Such sounds can be considered to be a part of product sound domain. Thus, understanding how product sounds are represented in the human mind would help interface designers or information ergonomists to design more intuitive user interface designs.

Experiment 1

An earlier study has shown that listeners may fail to access to the correct mental representation in memory because they have categorized some sounds on the bases of onomatopoeias, psychoacoustical and temporal descriptions (see Chapter 1). Moreover, despite the high occurrence of source descriptions, provided labels might not always be accurate. Therefore, Experiment 1 was conducted to determine listeners' ability to identify and label product sounds using a free labeling paradigm.

Procedure

Twenty-nine sounds were presented, each of which representing one of the six perceptual product sound categories. The sounds were recordings of various electrical domestic appliances in operation. They were either selected from various sound effect CDs, or recorded in house conditions by using a recording apparatus, Boss BR-532, with a Sennheiser e865 microphone with a frequency response of 40Hz - 20kHz and free-field sensitivity of 3mV/Pa. They were maximum five seconds long and were saved in a stereo format with a sampling rate of 44.1 kHz and 16 bits.

Eighteen students of Delft University of Technology (8 male and 10 female) participated. The mean age was 24.5. Their task was to identify the source of the sounds and to type the sound description on a computer screen. The sounds were presented using an especially designed software on a Macintosh PowerBook G4 computer via Sennheiser HD 477 headphones. The loudness levels were adjusted to a comfortable listening level for each sound. The participants were not allowed to change the sound levels during the experiment.

Results

The sound descriptions provided by the participants passed through an identification scoring. Similar to Marcell's study (2000), the responses that semantically matched with the actual name of the sound source were marked correct and scored as '1'. Incorrect responses were scored as '0'. Table 1 presents the mean proportion correct for each sound over participants. The mean proportion correct over all sounds is .29. In the table, digital alarm clock sound has the highest proportion correct (.93) followed by vacuum cleaner (.82), mechanical alarm clock (.61), microwave oven bell (.57), and coffee machine water pouring (.50) sounds. All the other sounds have

Sounds		**Experiment 1**	**Experiment 2**		
Groups	**Names**	**% Correct**	**Alternative Causes (Categories)**	**Causal Uncertainty**	**Familiarity Rating**
Air	Mixer	0.00	22	1.02	4.91
Air	Hairdryer	0.11	26	1.16	4.45
Air	Vacuum cleaner - hand	0.43	27	1.14	4.74
Air	Vacuum cleaner	0.82	18	0.91	4.71
Air	Washing machine	0.21	29	1.24	4.69
Air	Washing machine - centrifuge	0.04	31	1.26	4.17
Alarm	Alarm clock - digital	0.93	17	0.90	5.31
Alarm	Setting - MO	0.46	40	1.48	4.41
Alarm	Finish Bell - MO	0.57	23	1.09	5.01
Alarm	Finish Beep - MO	0.43	34	1.38	4.05
Cyclic	Computer	0.06	38	1.47	3.78
Cyclic	Microwave oven	0.00	27	1.32	3.75
Cyclic	Kitchen hood	0.22	40	1.51	3.69
Cyclic	Dishwasher	0.06	33	1.30	4.38
Cyclic	Tumble dryer	0.25	26	1.16	4.59
Impact	On/off switch - KH	0.14	55	1.69	4.03
Impact	Door closing - MO	0.29	42	1.48	4.35
Impact	Toaster	0.00	47	1.53	4.08
Impact	On/off switch - V	0.00	47	1.62	4.16
Impact	Door opening - WM	0.04	49	1.64	4.00
Liquid	Boiling - CM	0.46	45	1.55	4.27
Liquid	Brewing - CM	0.39	40	1.55	3.37
Liquid	Pouring water - CM	0.50	39	1.33	5.08
Mechanical	Citrus press	0.22	38	1.45	3.68
Mechanical	Blender	0.06	42	1.51	3.73
Mechanical	Shaver	0.11	30	1.23	4.06
Mechanical	Hair clippers	0.06	27	1.25	4.24
Mechanical	Toothbrush	0.22	42	1.55	3.69
Mechanical	Alarm clock - mechanical	0.61	16	0.93	4.56

Table 1. Twenty-nine sounds are presented with the mean proportion correct responses from Experiment 1 and with the categories of alternative causes, causal uncertainty values, and the familiarity rating from Experiment 2. ('MO' for microwave oven, 'KH' for kitchen hood, 'V' for ventilator, 'WM' for washing machine)

proportion correct scores below .50. Mixer, microwave oven, toaster, and ventilator on/off switch sounds have the lowest proportion correct.

The mean proportion correct for each sound group was analyzed with an ANOVA with sound categories as the within subjects factor (6 levels). Figure 1 presents the mean proportion correct for each product sound category over participants. According to the figure, alarm sounds have the highest proportion correct (.60) followed by liquid sounds (.45) and impact sounds have the lowest proportion correct (.09) followed by cyclic sounds (.13). A significant effect for sound categories was found, F(*5,135*) = 13.73, p<.001.

Participants' incorrect responses were analyzed to determine why listeners were not able to assign a correct label to a sound. It was observed that a participant very often used the label of another sound that has a similar spectral-temporal structure (e.g., 'shaver' instead of 'hair clippers'). It was also observed that incorrect labels and the target labels represent the sounds that are members of the same sound category (e.g., mechanical sounds).

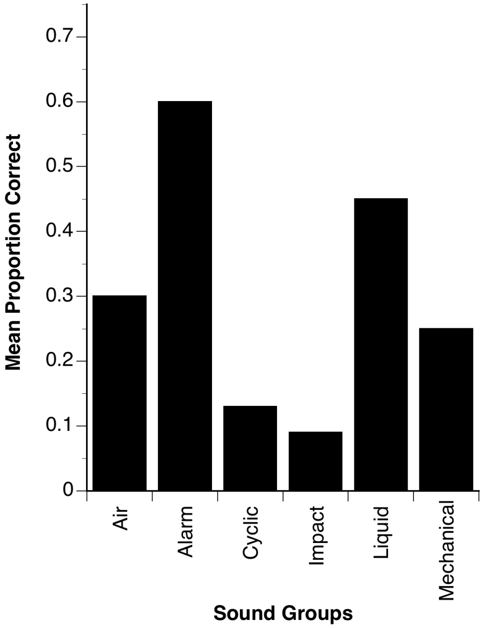


Figure 1. The mean proportion correct responses for each sound group over participants.

Conclusions

The results show that listeners fail to correctly label sounds caused by daily domestic appliances, except alarm sounds. The high scores for alarm sounds may be due to their distinct and structured spectral-temporal composition, because structured sounds are better represented in memory than unstructured or semi-structured sounds and retrieving the label of such sounds is easier (Özcan & van Egmond, 2007).

The results indicate that product sounds are not lexically well represented in memory. Naming errors occur during labeling, because listeners first fail to distinguish between sounds that belong to the same sound category. One of the reasons might be that listeners' insensitivity to the subtle differences in the structure of perceptually similar (noise-like) sounds. Moreover, fuzzy or incomplete encoding of the auditory information due to the noisiness in the structure of a sound may result in several mental representations, which further causes uncertainty in labeling. Therefore, Experiment 2 investigated how well product sounds are lexically represented in memory.

Experiment 2

Experiment 2 investigated the number and variety of lexical associations a product sound may have in semantic memory. In other words, this experiment was conducted to determine the causal uncertainty values for product sounds. One way to determine these values is by simply asking participants to provide the name(s) of any objects which they think are the causes of the sound. Obtaining the number of causes and determining the causal uncertainty values will allow us to understand whether the lexical impairment is due to multiple semantic representations.

Procedure

The same 29 sounds from Experiment 1 were used. Twenty-nine (2 male and 27 female) students of Plymouth University in UK participated. A participant's task was to identify all possible sources of the presented sound and write them down on a separate questionnaire sheet provided. The participants were explicitly encouraged to identify as many sounds as possible. For each name they provided, they rated their familiarity with the sound on a 7-point bi-polar scale (1-not familiar, 7-very familiar). The sounds were presented in a quite room through loudspeakers at a comfortable listening level.

Results

The distribution of the provided responses over participants showed that a participant provided maximum seven alternative causes for one sound. Of all the participants, 28% provided one, 33% two, 22% three, and 13% four alternative causes for one sound. A participant provided in average 2.3 alternative causes per sound. Table 1 presents the sum of the categories of alternative causes per sound. According to the table, participants agreed on minimum 16 (digital alarm clock sound) and maximum 55 (on/off switch sound of the kitchen hood) dissimilar categories of sound labels.

To determine the causal uncertainty values, entropy measures were obtained using Shannon's index for diversity (Zar, 1996). The same method was used in Ballas' studies (Ballas & Mullins, 1991; Ballas, 1993). Table 1 presents the causal uncertainty values for each sound—the lower the value, the higher the agreement between the participants. According to Table 1, digital alarm clock and the vacuum cleaner sounds have the lowest values (.90 and .91 respectively) followed by the mechanical alarm clock sound (.93). Moreover, the on/off switch sound of the kitchen hood had the highest value (1.69) followed by door opening sound of the washing machine (1.64) and on/off switch of the ventilator (1.62). In average, causal uncertainty values per sound group increased as follows: air (1.12), alarm (1.21), mechanical (1.32), cyclic (1.35), liquid (1.48), and impact (1.59).

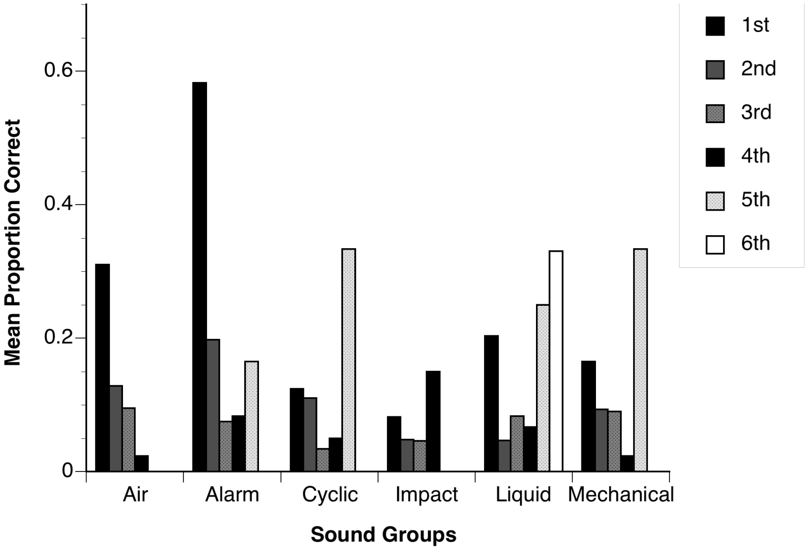


Figure 2. The mean proportion correct responses for each sound group over participants as a function of correct hit order.

It was also checked whether there were any correct hits among participants' responses. It was observed that 66% of the hits was in the first response, 21% was in the second, and 8% in the third. There were no hits in the seventh responses. Figure 2 presents the proportion correct for the product sound groups as a function of hit order. According to the figure, for alarm and air sounds, the first provided response was often correct; however, for cyclic, liquid, and mechanical sounds fifth (or sixth) responses have better hits. In addition, the mean proportion correct responses provided in the first attempt was correlated with the causal uncertainty values ($r = -.70$, $p < .001$, $N = 29$).

Table 1 also presents the familiarity ratings per sound. The average ratings per sound ranged between 3.37 (coffee brewing sound) and 5.31 (digital alarm clock sounds). The average familiarity rating for all sounds was 4.27 on a 7-point-scale. The familiarity ratings are correlated with the causal uncertainty values ($r = -.72$, $p < .001$, $N = 29$).

Table 2 presents the categories for the alternative causes given per sound and the numbers indicate the frequency of all responses for each category over all participants. In the table, the categories are presented in order of response frequency and sound names that were given only once overall participants were left out. It can be seen that the total number of similar alternative causes for one sound ranged from 53 (digital alarm clock sounds) through 93 (microwave oven finish bell).

Conclusions

The results confirm that product sounds have several lexical representations in memory because any given sound represents various objects/events that produce sound. These representations were mostly limited to within category similarities, although an across category similarity was observed between air and cyclic sounds. Thus, perceptual similarity between sounds is one of the reasons one of reasons for lexical impairment. Similar to Ballas' findings (1993), the results also show that identification accuracy decreases with high causal uncertainty. It is possible that the auditory information from product sounds is able to activate several semantic associations in memory at a time, thus confusions occur to pick the correct association and assign a label.

However, considering the low familiarity ratings (and their negative correlation to the causal uncertainty values), it is also possible that perceptual analysis of some sounds does not result in recognition; thus, no semantic or lexical association can be accessible, but the auditory information can still be mapped to the previously stored auditory representations in memory. Consequently, the result is a guessing strategy to find the best possible fit. The results of the hit order even confirm the guessing

Sound Group	**Sound name and frequency of responses**									
Air	**mixer**	**67**	**hair dryer**	**74**	**vacuum cleaner**	**58**	**washing machine**	**58**	**centrifuge**	**65**
	vacuum cleaner	26	hairdryer	15	vacuum cleaner	27	washing machine	16	vacuum cleaner	12
	hairdryer	9	vacuum cleaner	15	Hairdryer	5	dryer	6	airplane	11
	airplane	4	television	9	washing machine	4	dishwasher	4	hairdryer	9
	blender	4	radio	5	dryer	3	television	3	blender	2
	airplane engine	2	airplane	3	leaf blower	3	fan	2	drill	2
	plane starting	2	food mixer	3	airplane	2	food processor	2	juicer	2
	food mixer	2	airplane starting	2	carpet cleaner	2	radio	2	machinery	2
	lawn mower	2	dishwasher	2	lawn mower	2	toy car	2	television	2
Alarm	**alarm clock *d***	**53**	**setting *mo***	**67**	**finish bell *mo***	**93**	**finish beep *mo***	**63**		
	alarm clock	25	setting up mo	9	bell (microwave)	24	microwave	12		
	alarm	4	setting alarm	6	timer bell	16	alarm	6		
	timer	4	heart monitor	4	bell (for assistance)	12	fire alarm	3		
	bell (door)	3	phone keypads	4	bell (bicycle	5	timer	3		
	lorry reversing	3	microwave	4	clock	4	timer (oven)	3		
	fire alarm	2	alarm	2	triangle	4	beep (mo)	2		
	warning signal	2	beep (microwave)	2	xylophone	4	alarm being set	2		
			digital watch	2	bell	3	intercom	2		
Cyclic	**computer**	**60**	**microwave**	**55**	**kitchen hood**	**62**	**dishwasher**	**65**	**tumble dryer**	**56**
	airplane	10	dryer	7	dryer	6	washing machine	16	dryer	18
	air conditioner	4	washing machine	7	washing machine	5	dryer	8	washing machine	7
	wash. mach.	3	boiler room	3	boat engine	4	dishwasher	4	air conditioner	3
	airplane (inside	2	car	3	microwave	4	car	3	machinery	3
	boiler	2	dishwasher	3	air conditioner	3	air conditioner	2	car	2
	dishwasher	2	factory	3	car engine	3	boat engine	2	fan ass. oven	2
	fridge	2	air conditioner	2	extractor fan	2	extractor fan	2		
	heater	2	airplane	2	fan	2	rain	2		
	lift	2	fan	2	television	2	video camera	2		
Impact	**on/off switch *kh***	**72**	**door closing**	**77**	**toaster**	**79**	**on/off switch *v***	**64**	**door opening**	**66**
	door shutting	4	car door shutting	13	toaster	9	light switch	5	door shut. (met)	5
	hammering nail	4	car boot shut	5	hole puncher	7	hammering nail	4	door shutting	3
	stapler	3	door shutting	4	paper cutter	5	hitting wood	3	lid shutting (met.)	3
	switch (flicking	3	drum	4	stapler	5	chop. on board	2	toaster	3
	chopping food	2	dropping smth.	3	typewriter	5	knock on door	2	dropping smth.	2
	clock	2	someone falling)	3	spring	3	metronome	2	lid shutting	2
	dart hitting	2	stamp	3	eject button	2	nail gun	2	gun	2
	light switch	2	window shutting	3	let. box shutting	2	switch (flicking	2	lock going across	2
	metronome	2	boot shutting	2	scissors	2	tapping on wood	2	mo door shutting	2
	toaster	2	car door	2	stamp	2	ticking clock	2	nail gun	2
Liquid	**boiling *cm***	**77**	**brewing *cm***	**53**	**pouring water *cm***	**88**			‘	
	tap	6	toilet	4	water draining	16				
	water (boiling)	6	grinder	3	water pouring	15				
	water running	6	sucking a straw	3	water running	8				
	dishwasher	4	water draining	3	toilet	7				
	washing machine	4	coffee grinder	2	bath filling up	3				
	water draining	4	train	2	bath emptying	2				
	fish-tank pump	3	water pouring	2	filling up kettle	2				
	bath emptying	2			fountain	2				
	bath plug	2			stream	2				
	fountain	2			water pouring	2				
Mechanical	**citrus press**	**61**	**blender**	**68**	**shaver**	**63**	**hair clippers**	**66**	**toothbrush**	**66**
	blender	10	drill	7	shaver	18	buzzer (door)	13	shaver	5
	grinder	5	shaver	6	hair clippers	8	toothbrush	7	blender	4
	cement mixer	4	electric saw	5	toothbrush	4	hair clippers	6	drill	4
	food processor	3	television	4	buzzer	3	shaver	6	hedge cutter	4
	drill	2	buzzer (door)	2	buzzer (door)	2	buzzer	5	radio	4
	fire	2	e.sharpener	2	drill	2	drill	3	buzzer (door	2
	food mixer	2	hair clippers	2	electric saw	2	alarm	2	electric saw	2
	lawn mower	2	lawn mower	2	hedge cutter	2	electric saw	2	electricity	2
	microwave	2	roadwork	2			fluores. light	2	hair clippers	2

Note. ‘h’ for hand, ‘d’ for digital, ‘mo’ for microwave, ‘kh’ for kitchen hood, ‘v’ for ventilator, ‘wm’ for washing machine, ‘cm’ for coffee maker, ‘m’ for mechanical.

Table 2. Product sounds are presented with the categories of alternative causes from Experiment 2.

strategy. Product sounds that had low causal uncertainty values (e.g., air and alarm sounds) were identified in the first response; however, other sounds had better scores only in the fifth or sixth response. This also demonstrates that product sounds may be semantically represented in memory, but these representations for some sounds are fuzzy and not easily accessible.

Discussion

This study has provided insight into the variety of lexical associations that product sounds may have. It has been shown that listeners have difficulty in correctly identifying product sounds and that identification process for such sounds suffers from poorly represented auditory information both in the perceptual and lexical domains. The impairment in labeling mainly results from the attempt(s) to attribute meaning to not-recognizable auditory information. Thus, we can conclude that causal uncertainty, as commonly accepted, does not only result from multiple lexical associations that a sound may have in memory. This assumption is also supported by the high accuracy in identifying structured auditory information (e.g., alarm sounds) correctly and in the first attempt. Alarm sounds, for example, may have a relatively low causal uncertainty, yet they are associated with multiple concepts in memory. However, because of their structured spectral-temporal composition, it is easier to access the relevant semantic information. Therefore, the identification process depends on the perceptual analysis of the auditory information and cognitive processing benefits from the structure in spectral-temporal composition of a sound (Ballas, 1993; Deustch, 1980; Vanderveer, 1979).

Considering the cultural backgrounds of the participants (Dutch and English), one would expect that given identification responses would differ. However, within the participant responses, the authors have observed high similarities and have not encountered any cultural differences. This may be due to the similar life styles that people lead in both countries (e.g., using coffee-makers to prepare coffee, warming up food in a microwave oven or brushing teeth with electrical toothbrush). Thus, the results may be culture specific and represent the western European culture. Having said that, we predict that results will be similar in other countries (e.g., North-American) in which similar products are used to facilitate the modern life style.

An earlier study in visual cognition (Nickerson & Adams, 1979), which tested the visual memory for a daily object (namely, an American penny), has demonstrated that although people are able to correctly recognize a penny, they found it hard to reproduce its visual structure. Nickerson and Adams have concluded that the memory system stores ‘useful’ information. This is an interesting finding and may be adapted to the perception of product sounds. Many of the sounds are often used as use-cues to understand whether an appliance is working or functioning well. Except that alarm sounds are specially designed sounds to convey messages such as ‘food is ready’ or ‘wake-up’. For such sounds that are abstract (that do not derive from any natural event), semantic associations should be built instantly to code the exact meaning. However, for intrinsically occurring sounds (e.g., shaver sound) listeners may be reluctant to code their meaning. This might be because it is commonly assumed that a domestic appliance produces sound as a result of its functionality but not to convey a certain message. Thus, in the absence of a contextual situation it may be harder to recall the name of the sound of an appliance.

In this study, we have not checked the relationship between occurrence frequency of the sounds and their causal uncertainty values. Although these two factors for identifiability may be somewhat related, high occurrence frequency does not necessarily provide a faster and more accurate identification process (Ballas, 1993). For example, firing a gun is a rarely occurring event, and listeners are still able to identify the sound as good as they can identify the sound of a door bell. This indicates that there may be other factors that also influence the identifiability of environmental sounds. To speculate, emotional responses, the context in which the sound is presented, or familiarity may constitute other factors.

Next, we will investigate whether a provided context increases the identifiability of a product sound and decreases the ambiguity of causes by limiting the number of possibilities. If so, it will be investigated what type of context has a better influence in the identifiability of the product sounds. With this, we hope to provide more insight into other factors that may influence the identifiability of product sounds.

Moreover, the results of this study suggest that sound designers in the auditory display or product sound design field should consider that machinery sounds are not semantically or lexically well presented in memory. In addition, Özcan and van Egmond (2007) has suggested that visual or verbal labels help to retrieve the semantic information; thus, sound designers should remember to include, perhaps, verbal or visual labels in their product sound related communications. Sweller and Chandler (1991; 1994) and Tindall-Ford, Sweller, and Chandler (1997) have shown that dual-mode presentation (visual and auditory) of visual information reduces the cognitive load and increases the learnability of the instruction materials. Similarly, sounds in auditory displays are always designed in relation to a specific function (e.g., warning, feedback, etc.). For example, in the user interface design, visual buttons could support the auditory icons, or verbal labels could support the auditory warnings. Thus, to access better memory representations, designers should consider the necessity of the use of verbal/visual labels.

Acknowledgement

This research was supported by HFES Europe Chapter VRC Exchange Grant. We also would like to thank Judy Edworthy and Liz Hellier for their collaboration at Plymouth University, UK.

References

Aljishi S, T. (1991). Why does heating water in a kettle produce sound? *American Journal of Physics, 59*(7), 628-632.

Ballas, J. A. (1993). Common factors in the identification of an assortment of brief everyday sounds. *Journal of Experimental Psychology: Human Perception and Performance, 19*(2), 250-267.

Ballas, J. A., & Mullins, T. (1991). Effects of context on the identification of environmental sounds. *Human Performance, 4*(3), 1991-1219.

Bartlett, J. C. (1977). Remembering environmental sounds: Role of verbalization at input. *Memory & Cognition, 5*(4), 404-414.

Berglund, B., & Nilsson, M. E. (2003). Identification of sounds from traffic. *Perceptual Motor Skills*, *97*(3), 675-688.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review, 94*(2), 115-147.

Bismarck, G. (1974). Timbre of steady sounds: A factorial investigation of its verbal attributes. *Acustica, 30*, 146-159.

Björk, E. A. (1985). The Perceived quality of natural sounds. *Acustica, 57*(3), 185-188.

Bonebright, T. L (2001) *Perceptual structure of everyday sounds: A multidimensional scaling approach*. Proceedings of the *International Conference on Auditory Display, Finland, 7*, 73-78.

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound.* Massachusetts: MIT Press.

Cabe, P., & Pittenger, J. B. (2000). Human sensitivity to acoustic information from vessel filling. *Journal of Experimental Psychology: Human Perception and Performance, 26*(1), 313-324.

Chiu, C. Y. P., & Schacter, D. L. (1995). Auditory priming for nonverbal information: Implicit and explicit memory for environmental sounds. *Consciousness and Cognition, 4*(4), 440-458.

Coward, S. W., & Stevens, C. J. (2004). Extracting meaning from sound: Nomic mappings, everyday listening, and perceiving object size from frequency. *Psychological Record*, *54*(3), 349-364.

Crowder, R. G. (1993). Auditory memory. In A. S. Adams & E. Bigand (Eds.), In A. S. Adams & E. Bigand (Eds.), *Thinking in sound: The cognitive psychology of human audition*. New York: Oxford University Press.

Cummings, A., Ceponiene, R., Koyama, A., Saygin, A. P., Townsend, J., & Dick, F. (2006). Auditory semantic networks for words and natural sounds. *Brain Research, 1115*(1), 92-107.

Damasio, A. (1999). *The feeling of what happens: Body, emotion and the making of consciousness*: London: Vintage.

Deutsch, D. (1980). The processing of structured and unstructured tonal sequences. *Perception & Psychophysics, 28*(5), 381-389.

Edworthy, J., Hellier, E., & Hards, R. (1995). The semantic associations of acoustic parameters commonly used in the design of auditory information and warming signals. *Ergonomics, 38*(11), 2341-2361.

Fabiani, M., Kazmerski, V. A., Cycowicz, Y. M., & Friedman, D. (1996). Naming norms for brief environmental sounds: effects of age and dementia. *Psychophysiology, 33*(4), 462-475.

Gaver, W. W. (1989). The SonicFinder: An interface that uses auditory icons. *Human Computer Interaction, 4*, 67-94.

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

Gaver, W. W. (1993). *Synthesizing auditory icons.* Paper presented at the Interchi '93.

Guillaume, A. Pellieux, L. Chastres, V., Blancard, C., & Drake, C. (2005) *How long does it take to identify environmental sounds*. *Proceedings of the International Conference on Auditory Display, Australia*, *10*, 1-5.

Gygi, B., Kidd, G. R., & Watson, C. S. (2004). Spectral-temporal factors in the identification of environmental sounds. *Journal of the Acoustical Society of America, 115*(3), 1252-1265.

Handel, S. (1989). *Listening*. Massachusetts: MIT Press.

Hermes, D. J. (1998). *Auditory material perception* (Annual Progress Report No. 33): IPO.

Keller, P., & Stevens, C. (2004). Meaning from environmental sounds: Types of signal-referent relations and their effect on recognizing auditory icons. *Journal of Experimental Psychology: Applied, 10*(1), 3-12.

Kendall, R. A., & Carterette, E. C. (1993). Verbal attributes of simultaneous wind instrument timbres: I. von Bismarck's adjectives. *Music Perception, 10*(4), 445 - 468.

Kunkler-Peck, A. J., & Turvey, M. T. (2000). Hearing shape. *Journal of Experimental Psychology: Human Perception and Performance, 26*(1), 279 - 294.

Lass, N. J., Eastham, S. K., Parrish, W. C., Scherbick, K. A., & Ralph, D. M. (1982). Listeners' identification of environmental sounds. *Perceptual and Motor Skills*, *55*, 75-78.

Li, X., Logan, R. J., & Pastore, R. E., (1991). Perception of acoustic source characteristics: Walking sounds. *The Journal of Acoustical Society of America, 90*(6), 3036-3049.

Lutfi, R. A. (2001). Auditory detection of hollowness. *The Journal of Acoustical Society of America, 10*(2), 1010-1019.

Marcell, M. E., Borella, D., Greene, M., Kerr, E., & Rogers, S. (2000). Confrontation naming of environmental sounds. *Journal of Clinical and Experimental Neuropsychology, 22*(6), 830-864.

McAdams, S. (1993). Recognition of sound sources and events. In A. S. Adams & E. Bigand (Eds.), *Thinking in sound: The cognitive psychology of human audition* (pp. 146 - 198). New York: Oxford University Press.

Nickerson R. S., & Adams, M. J. (1979). Long-term memory for a common object. *Cognitive Psychology*, *11*, 287-307.

Özcan, E., & van Egmond, R. (2005). Characterizing descriptions of product sounds*. Proceedings of the International Conference on Auditory Display, Ireland, 11,* 55-60.

Özcan, E., & van Egmond, R. (2006). *Product sound design and application: An overview. Proceedings of the International Conference on Design and Emotion, Sweden*, *5,* 1-19.

Özcan, E., & van Egmond, R. (2007). Memory for product sounds: The effect of sound and label type. *Acta Psychologica, 126*(3), 196-215.

Özcan, E. & van Egmond, R., & Jacobs, J. (submitted) Bases for categorization and identification of product sounds.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology - Revue Canadienne De Psychologie, 45*(3), 255-287.

Solomon, L. N. (1958). Semantic approach to the perception of complex sounds. *Journal of the Acoustical Society of America, 30*(5), 421-425.

Stevenson, R. J., & Boakes, R. A. (2003). A mnemonic theory of odor perception. *Psychological Review, 110*(2), 340-364.

Sweller, J., & Chandler, P. (1991). Evidence for cognitive load theory. *Cognition and Instruction, 8*(4), 351-362.

Sweller, J., & Chandler, P. (1994). Why some material is difficult to learn. *Cognition and Instruction, 12*(3), 185-233.

Thompson, V. A., & Paivio, A. (1994). Memory for pictures and sounds: Independence of auditory and visual codes. *Canadian Journal of Experimental Psychology, 48*(3), 380-398.

Tindall-Ford, S., Chandler, P., & Sweller, J. (1997). When two sensory modes are better than one. *Journal of Experimental Psychology: Applied, 3*(4), 257-287.

Özcan, E., & van Egmond, R. (2006). *Product sound design and application: An overview. Proceedings of the International Conference on Design and Emotion, Sweden*, *5,* 1-19.

Vanderveer, N. J. (1979). Confusion errors in identification of environmental sounds. *Journal of the Acoustical Society of America, 65*, S60-S60.

Zar, J. H. (1996). Measures of dispersion and variability. *Biostatistical analysis*. New Jersey: Prentice Hall.

This chapter is based on the paper:

Özcan, E., & van Egmond, R. (2008). The effect of visual context of the identification of ambiguous environmental sounds. A revision submitted to *Acta Psychologica*.

Abstract

The influence of the specificity of the visual context on the identification of environmental sounds (i.e., product sounds) was investigated. Two different visual context types (i.e., scene and object contexts) that varied in the specificity of the semantic information and a control condition (meaningless images) were employed. A contextual priming paradigm was used. Identification accuracy and response times were determined in two context conditions and one control condition. The results suggest that visual context has a positive effect on sound identification. In addition, two types of product sounds (object-specific and event-specific sounds) were observed that exhibited different sensitivity to scene and object contexts. The results further suggest that conceptual interactions exist between an object and a context that do not share the same perceptual domain. Therefore, context should be regarded as a network of conceptually associated items in memory.

CHAPTER 5

⁂

THE EFFECT OF VISUAL CONTEXT ON THE IDENTIFICATION OF AMBIGUOUS PRODUCT SOUNDS

Imagine a room. A boiling pan is on the stove. The exhaust fan is running. On the counter lies an open pack of pasta, some vegetables with a knife, and other kitchen utensils. At this moment, a continuous and high-pitched sound is heard. The sound will probably be identified as a cooking timer, because the objects in this room refer to an everyday cooking activity. Now, imagine that a very similar sound is heard in the bedroom. Would it still be identified as a cooking timer, or as an old-fashioned ringing alarm clock? Probably the bedroom scene associates this sound with the alarm clock and excludes other causes of the sound (e.g., cooking timer, telephone, or school bell).

In everyday scenes, listeners seem to exhibit great accuracy in identifying otherwise ambiguous sounds. The production of environmental sounds is based on a non-arbitrary set of relationships between the sound source in action and the acoustic outcome of the event. In essence, listeners may be hearing, e.g., a metal bar continuously hitting a hollow metal object. However, it may be the context in which the sound is presented that determines the result of sound identification, i.e., the label of the sound. Furthermore, in isolated conditions auditory identification may suffer from the multiplicity of the activated sound sources (i.e., lexical associations), thus may not always result in accurate identification.

In this study we investigate whether a visual context influences auditory source identification using a contextual priming paradigm. In addition the specificity of visual information was manipulated and its effect on identification accuracy and response times were checked. In general, the results aim to provide insight into interactions between multiple perceptual systems (visual and auditory) and the semantic system. In the following sections we will discuss the effect of context on object identification by providing evidence from theories on object and scene perception, cross-modal audio-visual interactions, and auditory cognition.

Object identification in context

Short exposures and incomplete representations of visual information are sufficient to activate categorical or associative structures in memory (Antes, Penland, & Metzger, 1981; Bar, 2004; Bar & Aminoff, 2003; McCauley, Parmelee, Sperber, & Carr, 1980; Palmer, 1975). This activated network is regarded as context frames (Bar, 2004; Bar & Aminoff, 2003). Context frames are rich in information and can contain both perceptual and semantic knowledge. Rules and constraints exist that determine the relationship between the items of a context frame (Biederman, 1981; De Graef, Lauwereyns, & Verfaillie, 2000; Gordon, 2004). Therefore, semantic consistency of a context and a target object can be determined as fast as 80 ms (Davenport & Potter, 2004). In conclusion, object identification is not isolated from contextual scenes in which they likely occur (Bar, 2004; Biederman 1981; Heit & Barsalou, 1996; Palmer, 1975).

Three types of contextual effects have been discussed with respect to the stages of identification (Henderson & Hollingsworth, 1999). First, expectations derived from scene knowledge interact with the perceptual analysis of object (i.e. feature extraction and integration) (Biederman, 1981). That is, an activated context frame may sensitize the representation of all the context related objects. Then, when the visual object has been sufficiently analyzed, its most likely perceptual construction is mediated by the contextual activation (Bar, 2004; Bar & Aminoff, 2003). Secondly, context-object interactions occur at the matching stage, when perceptual descriptions are matched to long-term memory representations (Bar & Ulmman, 1996; Palmer 1975). That is, a rough perceptual analysis of the object is sufficient to derive the salient structural features. These features are then used to match with context-activated representations in long-term memory. Thirdly, object identification (including the matching stage) is isolated from context frames. That is, context frames have no facilitating effect on the perceptual processing (i.e., feature extraction and integration) of the target object (Hollingworth & Henderson, 1998). Both perceptual analysis and matching to long-term memory operate independent of context knowledge (Hollingworth & Henderson, 1999). Thus, unlike the interactions occurring on a pre-semantic phase, object (lexical) identification may be interfered by guessing strategies in later stages during cognitive processing (De Graef, Christiaens, & d'Ydewalle, 1990). Such effects may occur perhaps if both the scene and the objects are equally identifiable, but somewhat the relationships between the two (e.g., spatial, semantic) are violated.

Theories of object identification in context have concentrated on visual perception and semantic knowledge. Thus, it is not clear to what extent information from a visual context influences the sound identification. However, strong evidences have been found for cross-modal interactions that influence perceptual processes (see Spence & Driver, 2004, for an overview). For example, a unitary percept can result from audio-visual interactions at encoding for speech (McGurk & MacDonald, 1976), for abstract images and tones (Shimojo & Shams, 2001), and for synchronized audio-visual events (Ecker & Heller, 2005). However, such perceptual convergence may be a result from the lack of source identification. In addition, the auditory and visual events need to be coupled and presented as one event. For example, playing piano notes activates auditory codes congruent with the finger movements. This activation is stronger for experts than for novices and occurs prior to auditory identification at movement-to-sound matching phase (Hasegawa, et al., 2004). Similarly, Saldaña and Rosenblum (1993) investigated the congruency between (musical) hand gestures (cello bow or pluck) and an auditory continuum from bow to pluck. They eliminated the possibility of semantic interactions for the explanation of visual dominance on auditory judgments. They have suggested that the influence of visual information on auditory identification occurs more on an extra-modular cognitive level; and is not a true McGurk effect (i.e., a unitary percept). This finding brings to mind the possibility whether such extra-modular interaction lies in the holistic object information.

Latest theories on object representation in memory suggest that modality specific systems keep information about the relevant properties of an object (Barsalou 1999; Barsalou, 2008; Paivio, 1991). For example, visual information of an object is stored in the visual system and auditory information in the auditory system. Simultaneous encoding of multimodal information cause inter-system interactions that start as early as 40 ms (featural analysis) and continue within the 200 ms of pre-semantic process activating both modality specific and nonspecific areas (Giard & Peronnet, 1999). Paivio (1991; see also te Linde, 1984) has discussed that amodal conceptual information may be the link between the perceptual systems. If such a link is established between two perceptual systems at encoding, it can further influence recognition and recall performance (Edworthy & Hards, 1999; Özcan & van Egmond, 2007a; Paivio, 1991; Thompson & Paivio, 1994). The conceptual and perceptual representations of an object also have verbal correspondents which are separately stored in the semantic system. The object concept is directly represented by a label, which is a lexical representation and serve as a shortcut to access to the object's concept (Jescheniak, Hantsch, & Schriefers, 2005). Labeling at encoding also strengthens the conceptual network of an object. Although, retrieving auditory codes via labels is not very likely, exposure to an auditory stream may activate a lexicon and other semantically related objects (Bartlett, 1977; Chiu & Schacter, 1995; Stuart & Jones, 1995). Thus, there may be a one-way mnemonic link from a perceptual to a lexical store.

Environmental sound identification benefits from the additive effect of visual information because the main part of the auditory identification process is the determination of the sound's source and its relation to its environment (Kubovy & Valkenburg, 2000; Yost, 1991; see also Ballas, 1993; Fabiani, Kazmerski, Cycowicz, & Friedman, 1996; Marcell, Borella, Greene, Kerr, & Rogers, 2000; Vanderveer, 1979). Thus, when there is access to the concept and imagery of a sound source, an environmental is identified. Even if the identification is not required, a conceptual (and semantic) activation always occurs upon the perception of an environmental sound and prior to the access of lexical representations (Cummings, Ceponiene, Koyama, Saygin, Townsend, & Dick, 2006; Orgs, Lange, Dombrowski, & Heil, 2007). In general, conceptual network of environmental sounds may consist of basic, sub- and super-ordinate level concepts such as: locations (bathroom, kitchen), events (accident, sleep), objects (alarm clock, paper), situations (sickness, danger), or unidentified concepts (e.g., noise) and onomatopoeias (sound imitations such as bang, click).

The activation of the conceptual network may occur by either other auditory information or visual information that are both related to the object concept. Ballas and Mullins (1991) have explored this by presenting conceptually relevant auditory context to prime homonymous sounds (e.g., bacon frying and fuse burning). For example, food preparation sounds such as slicing and chopping would prime a bacon frying sound, but would not prime a fuse burning sound. However, because the context sounds were also somewhat ambiguous, the identification of the homonymous sounds did not improve. For positive contextual effects to occur, the items constituting context should have strong conceptual and lexical representations in memory. Perhaps, easy-to-identify everyday visual objects may have facilitated the auditory identification as the associations between the sound and context would be clearer.

Furthermore, the activation of the common conceptual network for object and context can be controlled. In memory, object concepts can bind different types of items that share a common attribute, and items that are linked to each other with varying degrees of associations (Medin, Lynch, & Solomon, 2000; Mervis & Rosch, 1981;

Rosch 1978; Tversky & Hemenway, 1984). For example, a concept can be a 'juicer', a 'kitchen', or just an 'orange'. Although these concepts may be overlapping on the knowledge they provide, they may still activate different objects that are specific to each concept in memory (De Wilde, Vanoverberghe, Storms, & De Boeck, 2003; Heit & Barsolou, 1996). For example, Bar and Aminoff (2003) have suggested that context-typical objects (e.g., roulette for casino) activate more rapidly semantic associations compared to objects that do not belong to any specific context (e.g., cherries). The activation may also depend on the level of conceptual associations (e.g., a super-ordinate level or basic level). Using pictures, Jolicoeur, Gluck, and Kosslyn (1984) have shown that objects are categorized faster when they represent basic level concepts rather than super-ordinate level concepts (see also Mervis & Rosch, 1981). Some associations elicit stronger activations for well-established category representations (Barsalou, 1983). For example, a kitchen concept may activate objects that constitute the kitchen scene such as pans, food processor, plates, table, sink, or other concepts such as cooking, eating, washing. In another example, an orange concept may activate other related objects such as knife, cutting, citrus press, cup, or other concepts such as eating, making juice.

This study

Literature suggests that contextual effects on identification are a result of perceptual or conceptual interactions between a weak item that needs to be identified and a strong item. A provided context may guide the perception of an object or may bias how the object is labeled. Therefore, especially the identification of an ambiguous object can be facilitated by the presence of contextual information. Positive effects of contextual information have often been found within the same perceptual system (e.g., visual object and a visual context) or between a perceptual system and a semantic system (e.g., visual object and a verbal attribute as context). Not many studies have reported explicitly the effect of context between different perceptual systems (e.g., auditory and visual).

Therefore, in this study, using a contextual priming paradigm, sound identification was tested within a *congruent context* (scene vs. object) and in a *control condition* in which no contextual cues were provided. Two types of images were presented as visual context: scenes and single objects. The object images were chosen to provide a more specific context compared to scene images. In addition, abstract images that contained meaningless pattern were used in the control condition.

The environmental sounds used in this study have a high occurrence frequency in everyday situations. These sounds could be conceptually associated to certain domestic scenes. Therefore, sounds produced by domestic appliances (i.e., product sounds) such as hairdryer, shaver, coffeemaker were used. The conceptual network for product sounds derive from the perceptual and cognitive judgments of the sounds (Özcan & van Egmond, 2007a; see also Chapter 1). Product sounds are mostly described by the product and the action causing the sound (shaver, door closing), the locations in which they occur (bathroom, work), abstract meanings (food is ready), the acoustic and temporal structure (sharp, continuous), etc. (Özcan & van Egmond, 2005). Because some product sounds evoke certain conceptual associations more often, six different categories can be distinguished within the domain of product sounds. These sound categories are: air, alarm, cyclic, impact, liquid, and mechanical sounds. The ambiguity measure (i.e., causal uncertainty) was calculated for each product sound in an earlier study (Özcan & van Egmond, 2007b; see the Appendix for values).

It was hypothesized that (a) identification accuracy for the sounds would be higher for the congruent context and lower for the control condition and (b) response times would be faster if the context specificity increased.

Experiment

Method

The experiment was a 3 x 6 mixed factorial design, with contextual situations (control condition, room and object contexts) and product sound categories (air, alarm, cyclic, impact, liquid, and mechanical) as factors. The experiment consisted of two phases: image identification and sound identification.

Participants

Sixty participants (35 male and 25 female), students and employees of industrial design engineering at Delft University of Technology, participated. The mean age was 24 years. Twenty participants were randomly assigned to each of the three experimental conditions formed by the context type. For control condition, 8 male and 12 female participants participated with the mean age of 23.6 years. For scene context condition, 12 male and 8 female participants participated with the mean age of 27.2 years. For object context condition, 15 male and 5 female students participated with the mean age of 21.1 years. All participants reported normal hearing and had normal or corrected-to-normal vision.

Stimuli

Table 1 summarizes the stimulus set that was used. In the table, short descriptions of the product sounds are given in terms of product types and the specific actions of the products, sound duration and the categories to which product sounds belong to. In addition, Table 1 presents the labels of the corresponding images (scene and object) that served as visual context for each sound. (Further details are given below.)

Auditory stimuli

Twenty-nine product sounds were used that represented the afore-mentioned six product sound categories (i.e., air, alarm, cyclic, impact, liquid, and mechanical sounds) (see Table 1). The sounds were taken from sound-effect CDs or were recorded using a recording apparatus (Boss BR-532, with a Sennheiser e865). Air and mechanical sounds were represented by six sounds, cyclic and impact sounds by five sounds, alarm sounds by four sounds, and liquid sounds by three sounds. The average duration of the product sounds varied between categories. For example, impact sounds have a short duration and air or mechanical have long duration. The duration of the sounds was limited in duration. Sounds longer than 5 s were trimmed to a duration of approximately 5 s (using Felt Tip Sound Studio v2.1). Sounds shorter than or equal to 5 seconds were not changed. Thus, the duration of the sounds varied between 302 ms and 5050 ms. The duration of air, cyclic, and mechanical sounds ranged between 4969 and 5050 ms (except for one mechanical sound which lasted 2171 ms). The duration of alarm sounds ranged between 1759 and 3100 ms. The duration of impact sounds ranged between 302 and 859 ms. The duration of liquid sounds ranged between 3936 and 4992 ms. The sounds were recorded at CD quality and were presented at a similar comfortable listening level preserving the natural variation in the loudness of sounds. The loudness levels ranged between 65 dB and 75 dB.

Visual stimuli

An abstract image with no conceptual relation was used as a control condition. Two types of visual context were chosen: object context and scene context (see Figure 1 and Table 1 for examples). The context images that have high conceptual relationship to the sound source information (i.e., products) were used. The scenes in which products most likely occur (toothbrush in a bathroom) and objects that are most likely used while interacting with a product (toothpaste and a toothbrush) are chosen to provide common conceptual associations. The specificity of this conceptual association should be higher for objects than for scenes. All images were saved in jpg format having a canvas of 590 x 470 pixels with 72 dpi in resolution, and presented in the center of a computer screen, at a distance of approximately 50 cm. Four judges agreed that photos represented the objects-to-be-identified well as context.

*Control Condition***.** Eight abstract images were chosen and randomly assigned to 29 product sounds. The images were chosen using only one guideline: the images were selected especially not to evoke any conceptual association with any contextual objects/situations/events. They were digital drawings with different colour traces.

Scene context. Five living environments were selected to represent the scene context images. Scene context images were colour photos that were taken by the authors. Two guidelines were used in the preparation of the room images: (1) the photos showed a part of a living environment (i.e., bathroom, bedroom, kitchen, living room, and office) where products are commonly used; (2) because the presence of domestic appliances could evoke some conceptual associations, no electrical appliances were shown in the photos.

Object context. Eighteen objects were selected to represent the object context images. An object essential in the product-user interaction was selected as a context. For example, an orange was selected to activate the conceptual association to a citrus press sound or a hairbrush was selected to a hairdryer sound (for examples, see Table 1 and Figure 1). Object context images were colour photos that were taken by the authors. The photos were taken using a proper angle that shows the identifiable features of the object (i.e., canonical perspective described by Palmer, Rosch, & Chase, 1981). Only the object was visible in a photo. All the objects had a grey background to reduce the contrast between figure and background, and to make the object more salient.

Apparatus

The stimuli were presented using a specially designed application developed using the Trolltech Qt (Mac OS X - free edition) tool kit. The application ran on a Macintosh Powerbook G4 1.33 GHz computer with 12" screen. The auditory stimuli were presented through AKG Studio Monitor K240DF 2x600 Ohm headphones.

An external button-box was designed to measure response time and the registration of yes-no answers using two buttons. question. The ‘no’ button registered the negative response. The button-box registered the response time in milliseconds after a button was pressed. The accuracy of the registration of the times was 1 ms. The button-box was connected to USB connection to the computer. An analogue connection triggered the internal clock (times) of the button-box to measure the response time. The timer was triggered with a 22 kHz pulse at 70 dB (SPL) with 50 ms in duration. The experiment took place in a quiet room.

Procedure

Each participant received a written explanation of the purpose of and the instructions for the study. A participant was seated in front of the screen at a distance of approximately 50 cm. The entire study consisted of two phases: (1) image identification and (2) sound identification.

Groups	Sounds	Duration (ms)	Scene Images	Object Images
Air	Hairdryer	5004	Bathroom	Hair brush
Air	Mixer	5027	Kitchen	Cake form
Air	Vacuum Cleaner	5004	Living room	Carpet
Air	Vacuum Cleaner (hand)	5016	Living room	Dirt and crumbles
Air	Washing Machine	5039	Bathroom	Laundry basket
Air	Washing Machine (centrifuge cycle)	5016	Bathroom	Laundry basket
Alarm	Alarm Clock (digital)	2450	Bedroom	Pillows
Alarm	Microwave Oven (beeps)	2868	Kitchen	Instant meal
Alarm	Microwave Oven (finish beep)	3100	Kitchen	Instant meal
Alarm	Microwave Oven (finish bell)	1759	Kitchen	Instant meal
Cyclic	Computer	5004	Office	Desk
Cyclic	Dishwasher	5027	Kitchen	Dirty dish
Cyclic	Kitchen extractor	4969	Kitchen	Pan
Cyclic	Microwave oven	4992	Kitchen	Instant meal
Cyclic	Tumble dryer	5004	Bathroom	Drying rack
Impact	Kitchen extractor (on-off button)	360	Kitchen	Pan
Impact	Microwave Oven (door closing)	584	Kitchen	Instant meal
Impact	Toaster	714	Kitchen	Bread slices
Impact	Ventilator (on-off button)	302	Living room	Cold drink
Impact	Washing Machine (door opening)	859	Bathroom	Laundry basket
Liquid	Coffee Maker (coffee brewing)	3936	Kitchen	Cup
Liquid	Coffee Maker (water boiling)	4992	Kitchen	Cup
Liquid	Coffee Maker (water pouring)	4528	Kitchen	Cup
Mechanical	Alarm Clock (mechanical)	2171	Bedroom	Pillows
Mechanical	Blender	4917	Kitchen	Milkshake
Mechanical	Citrus Press	5027	Kitchen	Orange (half)
Mechanical	Hair Clippers	5050	Bathroom	Comb
Mechanical	Shaver	5050	Bathroom	Bowtie
Mechanical	Toothbrush	5027	Bathroom	Toothpaste

Table 1. Product sounds, the sound category they belong to, their durations, and the descriptions of the type of context in which they were presented.

In the *image identification* phase, images corresponding to a specific condition were presented one at a time. A participant's task was to identify the image presented on the screen and type a brief description of the image. For the abstract images, participants were asked to type any description that fits the abstract images. A participant received no feedback during image identification and could proceed to the next phase after identifying all the images.

In the *sound identification* phase, sounds were presented one at a time with the images corresponding to the condition. However, a corresponding context image appeared on the screen minimum 500 and maximum 2000 ms before the sound (the image appearance duration was kept random to avoid conditioned automatic responses). An image stayed on the screen until the sound stopped playing. A participant's task was to identify the sound with the help of the image presented on the screen. A participant fulfilled the task by (1) pressing 'yes' or 'no' on the button-box as soon as possible depending on whether they could identify the sound, and (2) describing verbally (i.e., written) what (s)he had thought to have identified .

The entire experiment was self-paced and there were no pauses between the two phases of the experiment. The stimuli were always randomly presented in each phase and for each participant. Participants identified the context images and became familiar with them prior to the identification trials for sounds with context. This was done to avoid semantic interference for the identification of both context and the sound. Unfamiliar visual context could, for example, delay the response time because listeners would be processing the visual information instead of focusing on the auditory information. By doing this, we hoped to provide more consistent identification response times.

Control Condition Scene Condition Object Condition

Figure 1. Examples of images presented as control condition, scene, and object context.

Results

The responses for the image identification were analyzed separately from the sound identification data. Images were considered correctly identified when the descriptions semantically matched the target image name on a basic-level association for object images (e.g., a toothpaste) and on a super-ordinate-level associations for scene images (e.g., a bathroom). Correct responses were scored as '1' and incorrect responses as '0'. Any response for abstract images (control condition) was accepted as correct. The sums were divided by the number of images within a context type. Thus, proportion correct for scene images was .99 and for object images was .98. This confirmed that the chosen images were identifiable.

Two dependent variables for sound identification were measured: identification score in proportion correct and response time in milliseconds. Response time data were analyzed with an ANOVA with the following factors: context type (3 levels), sound type (6 levels), and identification type (3 levels). Identification data were analyzed with an ANOVA with the following factors: context type (3 levels) and sound type (6 levels).

Identification

Participants' identification responses were divided into three groups: 'yes' responses that were *correctly* identified; 'yes' responses that were *incorrectly* identified; and 'no' responses for *no* identification. Sounds were considered correctly identified if responses semantically matched the target sound names on a basic-level association (e.g., an electric toothbrush) and were considered incorrectly identified if responses semantically mismatched the target sound names. The sounds with 'yes-correct' responses were scored as '1', with 'yes-incorrect' responses as '-1', and 'no' as '0'. The sums were divided by the number of sounds within a sound category and within a context type. Table 2 presents the proportion correct for context type and for sound type. In the table, for the context type, the highest identification score is for the object context (.25) and the lowest is for the control condition (-.32). For the sound type, the highest identification score is for alarm sounds (.47) and the lowest is for cyclic sounds (-.27).

Figure 2 presents the mean proportion correct as a function of context and sound type. It can be seen in the figure that alarm sounds in the object context have the highest proportion correct (.66); and cyclic sounds in abstract condition have the lowest proportion correct (-.50). In addition, a possible interaction effect can also be found in the figure between context type and sound type. That is, alarm and air sounds were better identified in scene and object context than in control condition: impact, liquid, and mechanical sounds were better identified in object context than in

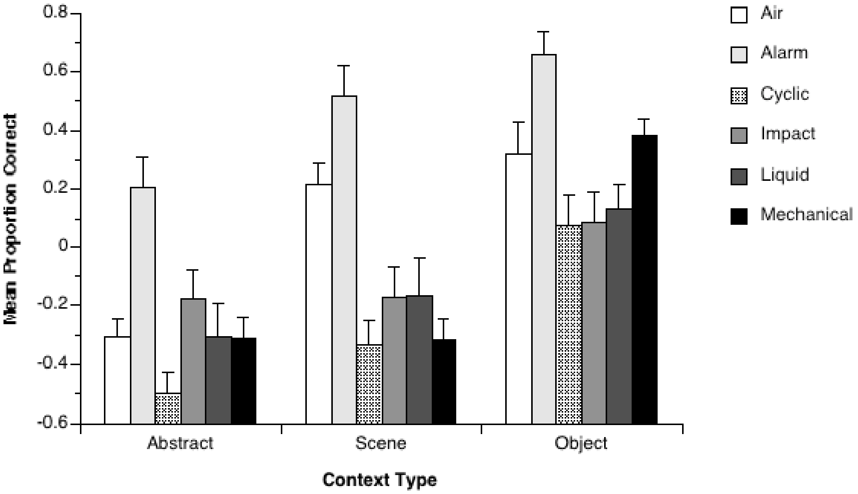


Figure 2. Mean proportion correct responses as a function of context type and sound type.

abstract and scene context. Significant main effects were observed for context type and for sound type, *F*(2, 342) = 62.14, *p* < .001 and *F*(5, 342) = 2.40, *p* < .05, respectively. The above-described interaction effect between context and sound type was confirmed, *F*(10, 342) = 2.40, *p*<.05. A post-hoc analysis was conducted to reveal which levels of the main effects differed significantly. For context type, all context types differed significantly from each other (*p*<.001). For sound type, alarm sounds differed significantly from all other sound types (*p*<.001); cyclic sounds differed significantly from alarm (*p*<.001) and mechanical sounds (*p*<.05).

Response Time

Table 2 presents the mean response times for context type, sound type, and identification type. In the table, within the context type, the response time for the object context is the fastest (3621.30 ms) and for the control condition is the slowest (4231.39 ms). Within the sound type, the average response time for alarm sounds is the fastest (3583.34 ms) and for the cyclic sounds is the slowest (4267.55 ms). Within the identification type, response time for the yes-correct identification is the fastest (3147.73 ms) and for no-identification is the slowest (4668.89 ms). In addition, response times for each product sound within a context condition (control condition, scene context, and object context) were correlated with their sound durations. The correlations between sound duration and reaction times were calculated for each context condition (abstract condition *r*=.81, object condition *r*=.54, and scene condition *r*=.55). In addition, the difference between the mean response time and the duration of the sound was calculated (control condition, *M* = 366.29 ms, *SD* =

1156.16ms; object condition, *M* = -428.12ms, *SD* = 1509.85 ms; scene condition, *M* = -321.42ms, *SD* = 1516.98 ms). Thus, if the context information was less specific, the reaction time became longer because people needed to listen longer.

Figure 3 presents the mean response times as a function of context and identification type. According to the figure, the reaction for the yes-correct identification in object condition is the fastest (2825 ms) and for the no-identification in the scene context the slowest (4837 ms).

Condition	Condition Type	Mean Proportion Correct	Mean Response Time (milliseconds)
Context	Control	-0.32 *(0.04)*	4164.80 *(108.23)*
	Scene	-0.08 *(0.05)*	3595.63 *(108.77)*
	Object	0.25 *(0.04)*	3519.28 *(102.19)*
Sound	Air	-0.12 *(0.06)*	4066.91 *(141.06)*
	Alarm	0.48 *(0.06)*	3279.80 *(142.32)*
	Cyclic	-0.26 *(0.06)*	4232.89 *(147.88)*
	Impact	-0.18 *(0.06)*	3082.01 *(125.36)*
	Liquid	-0.17 *(0.07)*	4259.83 *(162.76)*
	Mechanical	-0.01 *(0.07)*	3556.32 *(156.15)*
Identification	Yes (Correct)	-	3044.19 *(85.90)*
	Yes (Incorrect)	-	3839.32 *(89.60)*
	No	-	4551.26 *(133.67)*

Note. Italic numbers in parenthesis are the standard error of the mean.

Table 2. Mean proportion correct as a function of context and sound type, and mean response times in milliseconds as a function of context, sound, and identification type.

Figure 4 presents the mean response times as a function of identification type and sound type. According to the figure, the response time for the yes-correct identification for mechanical sounds is the fastest (2370 ms) and for the no-identification for air sounds the slowest (5133 ms). Significant main effects were observed for context type, sound type, and identification type, $F(2, 743) = 10.22$, $p<.001$, $F(5, 743) = 12.37$, $p<.001$, and $F(2, 743) = 53.11$, $p<.001$, respectively. Interaction effects were found for the context type and identification type, and for sound type and identification type, $F(4, 743) = 2.59$, $p<.05$, and $F(10, 743) = 2.90$, $p<.05$, respectively. A post-hoc analysis was conducted to reveal which levels of the main effects differed significantly. For context type, only control condition differed significantly from the other context types (*p*<.05). For sound type, alarm sounds differed significantly from air, cyclic, and liquid sounds (*p*<.001); impact sounds differed significantly from air, cyclic, and liquid sounds (*p*<.001). The response times for the object and the scene conditions for the yes-responses were significantly faster than for the abstract condition (*p*<.001).

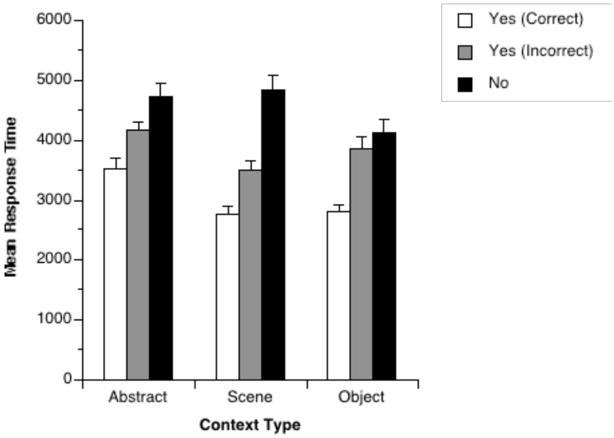


Figure 3. Mean response times as a function of context type and identification type.

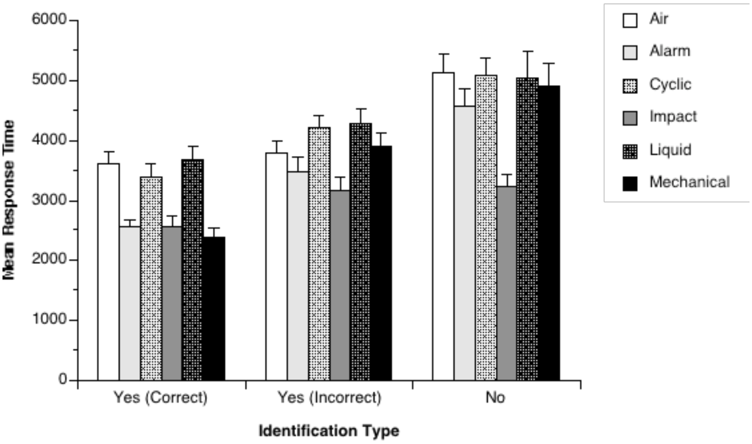


Figure 4. Mean response times as a function of identification type and sound type.

Discussion

Our main finding is that visual context positively affects the identification performance for ambiguous environmental sounds (i.e., product sounds). As predicted, this effect has been found for the identification accuracy as well as for the response time taken to identify a sound. The main trend is that the identification accuracy improves along with context specificity, however the duration of the response times benefit only from context not from context specificity.

Our second finding is that the identification performance differs depending on the sound type in an available context (scene or object). Although the identification accuracy in the object context has improved for all product sounds, the scene context only improved the recognition for air and alarm sounds. Interestingly, impact, mechanical, and liquid sounds do not benefit from scene context, but from object context. Only for cyclic sounds the effect of context specificity is as predicted. Although, the response times for correct identification decreased with available congruent context, no significant differences were found between the scene and object contexts. No-responses are faster in the object context. Conversely, yes-incorrect responses are faster for the scene context than for the object context.

The results suggest that the effect of context specificity is sound type dependent. The processing time of the auditory identification appears to be independent of the context specificity. That is, both context types equally facilitate the accurate auditory identification. Only no-responses benefit from context specificity. Thus the identification of some product sounds benefit from both scene and object context and another group only from object context. In the following sections these findings and the possible reasons will be discussed.

Context specificity and sound type

The associations that objects may have with each other or with a scene have been discussed extensively in literature. Barsalou (1983) focuses more on instances of categories at different levels, whereas Bar (2004) suggests that *context frames* represent the typical arrangements of objects in our environment. That is, a context frame provides basic global information and expectation-based shortcuts that facilitate the featural extraction and recognition of objects / relations that are sufficiently characteristics of context. A context frame and its specific objects actually pass through the same cortical analysis and activate the same cortical area (Bar & Aminoff, 2003). Both approaches suggest that strong associations to a context frame or typicality of the instances within a category are required for a contextual effect to occur.

Although these findings are mostly based on visual information processing theories, they may still explain why certain sound categories benefit from scene context and other categories from an object context. According to the conceptual network for product sounds, each product sound category can be characterized by certain basic concepts (see Chapter 1; Özcan & van Egmond, 2005). Some sounds (air and cyclic sounds) are characterized by the locations in which they occur, whereas others (impact, liquid, and mechanical) are characterized by interactions that cause the sound event (closing doors, rotating buttons). Alarm sounds are characterized by meaning (e.g., wake-up, food is ready) and are implicitly scene related.

Apparently, some product sounds are location specific (including alarm sounds because they are context-dependent auditory objects) and others are event / interaction specific. Location specific sounds may be an integral part of the scene in which they occur because the sound source (i.e., product) typically belongs to that scene (dishwashers in kitchens and washing machines in bathrooms). Event specific sounds may not have such a strong typical relation to the location because they incidentally occur in a scene and are caused by the parts of the product. For example, an on-off button can be considered as an atypical sound for any specific scene. Thus, their conceptual relation to a scene is not strong. Thus, object context may be activating relevant expectations regarding a sound event and the agents of the event. For example, a half orange may activate the act of orange squeezing, an empty coffee cup may activate pouring coffee, or dirty-clothes basket may activate putting clothes in washing machine.

In conclusion, context specificity may not always be a necessary factor that disambiguates the sound identification. Disambiguation may be facilitated with strong conceptual associations. For location-specific sounds this may be a scene, for event-specific sounds this may be an object.

The effect of context

As discussed above, a context frame has the capacity to activate various objects and events that are conceptually related to each other. This activation does not only occur within the semantic system but also in the perceptual systems (Bar, 2004). Therefore, we can assume that the conceptual identification of a context frame activates the perceptual composition of the context in terms of its visual, auditory, olfactory, or motor-behaviour representations, objects that semantically match the context and events that are likely to occur in the context.

We suggest that a conceptual network activated both by context frames and sounds may have played an essential role in the facilitation of the identification of auditory events. The sensory representation of the sound is established through a perceptual analysis. However, the result of this analysis may be prone to top-down influences. At a later stage, this representation will activate conceptual and semantic associations in long-term memory. Finally, a lexicon is accessed and a label (correct or incorrect) is assigned to a sound. Similarly, a visual context is perceptually analyzed. Accessing to a concept or a category that represents the visual percept results in identification and consequently the activation of the related items. These items may be of sensory and of semantic nature and they may also be restricted by the task (i.e., sound identification within context).

For a contextual effect to occur there needs to be an overlap between the representations activated by context and by the weak object-to-be-identified. First, contextual information may bias the perceptual process through which the structural content of the object is determined. Consequently, the encoding of the object will bias a context-activated perceptual representation. Second, context facilitates the process of matching the structural features of the object to those in long-term memory. At this stage, context-activated perceptual representations will determine the amount of structure that is required to recognize an object (Bar & Ullman, 1996; Biederman, 1981; Palmer, 1975). The third type of context effect occurs in post-perceptual stages such as matching to a lexicon (De Graef et al., 1990). Therefore, it is plausible that additional semantic knowledge facilitates the naming of an object.

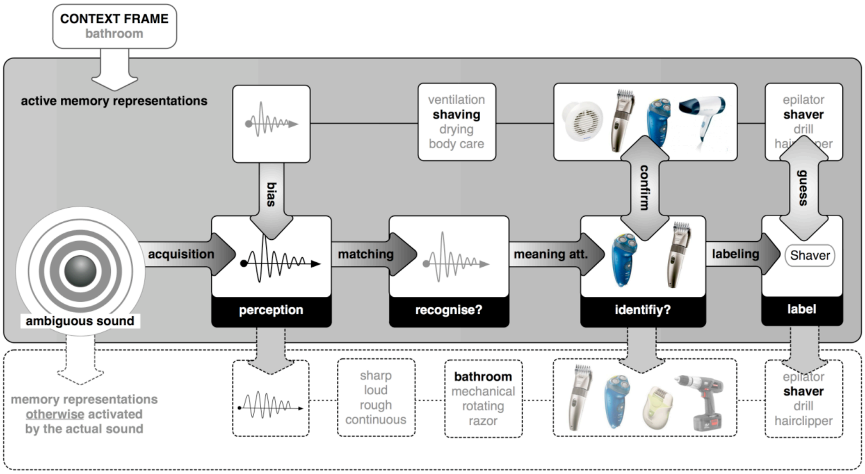


Figure 5. The effect of visual context on the identification of product sounds. The figure presents these effects as a function of the stages of a hypothetical identification process of an ambiguous product sound.

There may be two possibilities for the contextual effects of this study (see Figure 5). First, context frames may have intervened the auditory featural analysis at the pre-semantic stage. However, the audio-visual interaction probably did not result in a unitary percept. In our experimental set-up at the encoding of the sound the sound source was absent, thus a visual event was not coupled to an auditory event (e.g., Saldaña & Rosenblum, 1993). Alternatively, auditory codes may have been activated via context frames if auditory information is an integral property of the context frame (e.g., scene context). Second, context frames may have limited the activation of the sound concepts and labels at the post-perceptual stages. In the following paragraphs, the contextual effects within scene context and object context will be discussed in relation to location-specific and event-specific sounds.

Scene context

Our results suggest that the identification of location-specific product sounds (alarm, air, and cyclic sounds) benefit equally from a scene context as from an object context. This is an interesting finding because a scene is a concept at the super-ordinate level and would narrow down the possible lexical associations to only scene-specific ones, but not to the object-specific ones. Thus, a strong conceptual link may exist between these sounds and locations. One possibility is that audio-visual contextual interactions have occurred on a semantic level. A location-specific sound probably activates not only possible labels for the sound (sound source) but also activates knowledge where it can be heard. Consequently, a scene confirms plausibility of the sound. Alternatively, a scene context may have quickly activated the auditory representation of the scene, in which the target sound may be present. Such activation may bias the perceptual analysis of the ambiguous sound or facilitate the matching process to the long-term memory representations activated by the context frame.

It seems that scene context has not offered much facilitation for event-specific sounds. Incidental events and their properties may be atypical to a scene. Therefore, establishing a common conceptual network becomes difficult. If the event-specific sounds were identified as products (but not as events caused by product parts), then their semantic relation to the context frame would have been clearer.

Object context

Object context has provided equal facilitation for both location specific and event specific sounds. For both location-specific and event-specific sounds, one strong possibility is that the effect may have occurred on a semantic level. The object and the sound are mostly contextually related—they co-occur in certain situations. In addition, it seems very unlikely that the object context to activates the auditory codes of the target sound because the presented object is indirectly related to the sound but contextually related to the sound source. Consequently, this can only occur via semantic associations.

Furthermore, the positive effects of visual context on sound identification may have also been provided by the auditory information. Studies on auditory perception suggest that listeners are able to identify the properties of the sound event such as action, interacting materials, shape of the objects, (Cabe & Pittenger, 2000; Hermes, 1998; Kunkler-Peck & Turvey, 2000). Consequently, a sound that is identified as an interaction event may as well activate an object concept.

General conclusions

This study has demonstrated the positive effect of visual context on sound identification performance. This effect seems to be dependent on the sound type (location- and event-specific) as much as on the visual context type (scene and context). The visual context seems to facilitate mostly the post-perceptual processes and constrain mainly the semantic activation and guessing strategies. Nevertheless, it seems plausible that audio-visual interactions influence the auditory perceptual process if the visual context and the sound activates the conceptual network equally strong. Therefore, these contextual effects require a two-way conceptual interaction between the perceptual systems. This results in a rejection or confirmation whether the heard sound belongs to a visual context.

This study has provided evidence that audio-visual interactions occur for the identification of multi-sensory objects and that perceptual systems (in this case auditory and visual) are interconnected through conceptual associations. The findings suggest that *context* should be regarded as a network of associated items in memory that facilitates cognitive functions related to object identification. Conceptual associations by which several perceptual systems and a semantic system are interconnected constitute this network.

Limitations of this study

The auditory stimuli used in this study may not have been optimal for a couple of reasons. For example, not all sound categories contained the equal number of sounds (compare three liquid sounds and five mechanical sounds). The duration of the sounds also varied depending on the sound type (compare the short impact sounds and long liquid sounds). Especially, the differences in duration may make it difficult to compare response time results for different sound types. However, previous research (Chapter 1; Özcan & van Egmond, 2005, 2007a, 2007b) investigated the conceptual network of product sounds by using these sounds.

Therefore, using the same sounds helped to further understand the audio-visual interactions.

The visual stimuli set were not previously checked systematically on their association strengths to the sounds used. For example, the image of an orange may have activated the object ‘citrus press’ rather easily compared to the cold drink used for activating ventilator on/off button. This may have influenced the results, as strong conceptual associations may have not been activated with equal strength for all sounds. This may be critical as ideally visual context would constrain the identification by activating the sound source related concepts or associations, which would correspond to the sound activated concepts.

Future studies

Future studies could address to the afore-mentioned limitations of this study. The systematic arrangement of the association strength of the context frames in relation to the employed sounds is of importance. Furthermore, the effect of visual context on ambiguous sound identification should also be investigated using inconsistent contexts in order to see whether perceptual biasing effects would occur.

Regarding the semantic effects, verbal / written labels (e.g., text) could also be used as context (instead of visual context) to investigate whether the identification accuracy or response time would change and the semantic interaction effect will remain. Providing text could result in a faster access to semantic or conceptual associations in memory compared to visual contexts, because the semantic association will be more direct for a verbal description (see Özcan & van Egmond, 2007).

Acknowledgements

We would like to thank Suzanne Ringeling, Floor de Bruijn for running of the experiments, and Alexandre Vieira for voluntarily programming the experimental software in his free time. We are also very grateful to Ans Koenderink-van Doorn for her valuable comments on the design of the study and on the text.

References

Antes, J. R., Penland, J. G., & Metzger, R. L. (1981). Processing global information in briefly presented pictures. *Psychological Research, 43*(3), 277-292.

Ballas, J. A. (1993). Common factors in the identification of an assortment of brief everyday sounds. *Journal of Experimental Psychology: Human Perception and Performance, 19*(2), 250-267.

Ballas, J. A., & Mullins, T. (1991). Effects of context on the identification of environmental sounds. *Human Performance, 4*(3), 1991-1219.

Bar, M. (2004). Visual objects in context. *Nature Reviews: Neuroscience, 5*(8), 617-629.

Bar, M., & Aminoff, E. (2003). Cortical analysis of visual context. *Neuron, 38*(2), 347-358.

Bar, M., & Ullman, S. (1996). Spatial context in recognition. *Perception, 25*(3), 343-352.

Barsalou, L.W. (2008). Grounded cognition. *Annual Review of Psychology, 59*, 617-645

Barsalou, L. W. (1983). Ad hoc categories. *Memory & Cognition, 11*(3), 211-227.

Bartlett, J. C. (1977). Remembering environmental sounds: Role of verbalization at input. *Memory & Cognition, 5*(4), 404-414.

Biederman, I. (1972). Perceiving real-world scenes. *Science, 177*, 77-80.

Biederman, I. (1981). On the semantics of a glance at a scene. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual organization* (pp. 213-263). New York: Lawrence Erlbaum.

Cabe, P., & Pittenger, J. B. (2000). Human sensitivity to acoustic information from vessel filling. *Journal of Experimental Psychology: Human Perception and Performance, 26*(1), 313-324.

Chiu, C. Y. P., & Schacter, D. L. (1995). Auditory priming for nonverbal information: Implicit and explicit memory for environmental sounds. *Consciousness and Cognition, 4*(4), 440-458.

Cummings, A., Ceponiene, R., Koyama, A., Saygin, A. P., Townsend, J., & Dick, F. (2006). Auditory semantic networks for words and natural sounds. *Brain Research, 1115*(1), 92-107.

Davenport, J. L., & Potter, M. C. (2004). Scene consistency in object and background perception. *Psychological Science, 15*(8), 559-564.

De Graef, P., Christiaens, D., & d'Ydewalle, G. (1990). Perceptual effects of scene context on object identification. *Psychological Research, 52*(4), 317-329.

De Graef, P., Lauwereyns, J., & Verfaillie, K. (2000). *Attentional orienting and scene semantics* (Psych. Rep. Nr. 268). Leuven: University of Leuven, Laboratory of Experimental Psychology.

De Wilde, E., Vanoverberghe, V., Storms, G., & De Boeck, P. (2003). The instantiation principle re-evaluated. *Memory, 11*(6), 533-548.

Ecker, A. J., & Heller, L. M. (2005). Auditory-visual interactions in the perception of a ball's path. *Perception, 34*(1), 59-75.

Edworthy, J., & Hards, R. (1999). Learning auditory warnings: The effects of sound type, verbal labeling and imagery on the identification of alarm sounds. *International Journal of Industrial Ergonomics, 24*, 603-618.

Fabiani, M., Kazmerski, V. A., Cycowicz, Y. M., & Friedman, D. (1996). Naming norms for brief environmental sounds: effects of age and dementia. *Psychophysiology, 33*(4), 462-475.

Giard, M. H., & Peronnet, F. (1999). Auditory-visual integration during multimodal object recognition in humans: A behavioral and electrophysiological study. *Journal of Cognitive Neuroscience, 11*(5), 473-490.

Gordon, R. D. (2004). Attentional allocation during the perception of scenes. *Journal of Experimental Psychology: Human Perception and Performance, 30*(4), 760-777.

Gygi, B., Kidd, G. R., & Watson, C. S. (2004). Spectral-temporal factors in the identification of environmental sounds. *Journal of the Acoustical Society of America, 115*(3), 1252-1265.

Hasegawa, T., Matsuki, K., Ueno, T., Maeda, Y., Matsue, Y., Konishi, Y., et al. (2004). Learned audio-visual cross-modal associations in observed piano playing activate the left planum temporale. An fMRI study.

Cognitive Brain Research, 20(3), 510-518.

Heit, E., & Barsalou, L. W. (1996). The instantiation principle in natural categories. *Memory, 4*(4), 413-451.

Henderson, J. M., & Hollingworth, A. (1999). High-level scene perception. *Annual Review of Psychology, 50*, 243-271.

Hermes, D. J. (1998). *Auditory Material Perception* (Annual Progress Report No. 33): IPO.

Herz, R. S. (2003). The effect of verbal context on olfactory perception. *Journal of Experimental Psychology: General, 132*(4), 595-606.

Hollingworth, A., & Henderson, J. M. (1998). Does consistent scene context facilitate object perception? *Journal of Experimental Psychology: General, 127*(4), 398-415.

Hollingworth, A., & Henderson, J. M. (1999). Object identification is isolated from scene semantic constraint: evidence from object type and token discrimination. *Acta Psychologica, 102*(2-3), 319-343.

Jescheniak, J. D., Hantsch, A., & Schriefers, H. (2005). Context effects on lexical choice and lexical activation. *Journal of Experimental Psychology: Learning, Memory, & Cognition, 31*(5), 905-920.

Jolicoeur, P., Gluck, M. A., & Kosslyn, S. M. (1984). Pictures and names: making the connection. *Cognitive Psychology, 16*(2), 243-275.

Kubovy, M., & Van Valkenburg, D. (2001). Auditory and visual objects. *Cognition, 80*(1-2), 97-126.

Kunkler-Peck, A. J., & Turvey, M. T. (2000). Hearing shape. *Journal of Experimental Psychology: Human Perception and Performance, 26*(1), 279-294.

Marcell, M. E., Borella, D., Greene, M., Kerr, E., & Rogers, S. (2000). Confrontation naming of environmental sounds. *Journal of Clinical and Experimental Neuropsychology, 22*(6), 830-864.

McCauley, C., Parmelee, C. M., Sperber, R. D., & Carr, T. H. (1980). Early extraction of meaning from pictures and its relation to conscious identification. *Journal of Experimental Psychology: Human Perception and Performance, 6*(2), 265-276.

McGurk, H., & MacDonald, J. (1976). Hearing lips and seeing voices. *Nature, 264*(5588), 746-748.

Medin, D. L., Lynch, E. B., & Solomon, K. O. (2000). Are there kinds of concepts? *Annual Review of Psychology, 51*, 121-147.

Mervis, C. B., & Rosch, E. (1981). Categorization of natural objects. *Annual Review of Psychology, 32*, 89-115.

Orgs, G., Lange, K., Dombrowski, J. H., & Heil, M. (2006). Conceptual priming for environmental sounds and words: an ERP study. *Brain and Cognition, 62*(3), 267-272.

Özcan, E., & van Egmond, R. (2005). Characterizing descriptions of product sounds. *Proceedings of the International Conference on Auditory Display, Ireland, 11*, 55-60.

Özcan, E., & van Egmond, R. (2007a). Memory for product sounds: The effect of sound and label type. *Acta Psychologica, 126*(3), 196-215.

Özcan, E., & van Egmond, R. (2007b). How well do we identify product sounds? *Proceedings of the International Conference on Auditory Display, Canada, 13*, 234-241.

Özcan, E., van Egmond, R., & Jacobs, J. (submitted). Bases for categorization and identification of product sounds.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology - Revue Canadienne De Psychologie, 45*(3), 255-287.

Palmer, S. E. (1975). The effects of contextual scenes on the identification of objects. *Memory & Cognition, 3*(5), 519-526.

Palmer, S., Rosch, E. & Chase, P. (1981). Canonical perspective and the perception of objects. In J. Long & A. Baddeley (Eds.), *Attention and Performance IX* (p.p. 135 – 151). Hillsdale: Lawrence Erlbaum Associates.

Rosch, E. H. (1978). Principles of categorization. In E. H. Rosch & B. B. Lloyd (Eds.), *Cognition and Categorization* (pp. 27-48). Hillsdale, NJ: Lawrence Erlbaum Associates.

Saldaña, H. M., & Rosenblum, L. D. (1993). Visual influences on auditory pluck and bow judgments. *Perception & Psychophysics, 54*(3), 406-416.

Shimojo, S., & Shams, L. (2001). Sensory modalities are not separate modalities: Plasticity and interactions. *Current Opinion in Neurobiology, 11*(4), 505-509.

Spence, C., & Driver, J. (Eds.). (2004). *Crossmodal space and crossmodal attention.* Oxford: Oxford University Press.

Stuart, G. P., & Jones, D. M. (1995). Priming the identification of environmental sounds. *Quarterly Journal of Experimental Psychology, 48A*(3), 741-761.

Thompson, V. A., & Paivio, A. (1994). Memory for pictures and sounds: Independence of auditory and visual codes. *Canadian Journal of Experimental Psychology, 48*(3), 380-398.

Tversky, B., & Hemenway, K. (1984). Objects, parts, and categories. *Journal of Experimental Psychology: General, 113*(2), 169-197.

Vanderveer, N. J. (1979). Confusion errors in identification of environmental sounds. *Journal of the Acoustical Society of America, 65*, S60.

Yost, W. A. (1991). Auditory image perception and analysis: The basis for hearing. *Hearing Research, 56*(1-2), 8-18.

CONCLUSIONS

⁂

PART A

This part of the thesis started with the main question: how do people attribute meaning to product sounds? This question led into investigations into the identification of product sounds and meaningful associations that occur at different stages of the identification process. Accordingly, nine experiments were conducted that concentrated on perceptual processes (sound encoding, recognition) and cognitive processes (memory, categorization, labeling) that constitute the identification process. Furthermore, it was investigated whether sound identification

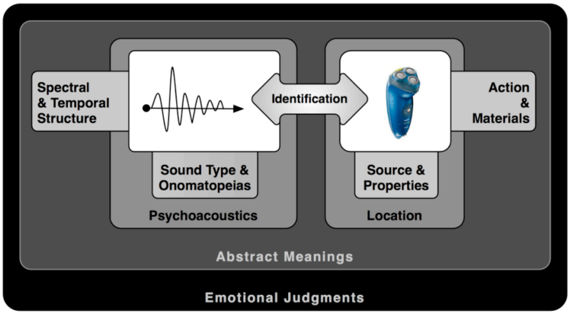


Figure 1. Conceptual network of product sounds.

occurs only within the auditory system or that other systems, especially the visual system, also contribute to this identification process. It was aimed to understand when and how different sound descriptions occur in this process and whether these sound descriptions are a result of an incomplete sound identification. How sound identification could benefit from additional semantic information was investigated.

The main conclusion is that the identification of a product sound is not restricted to the determination of the sound source. A product sound can evoke a large conceptual network that is composed of sound and source related information (see Figure 1). In addition, sound and source information are conceptually related, but not necessarily via the visual system. The visual system's contribution to sound identification occurs more on an extra-modular level. That is, a product concept consists of information regarding the visual, auditory and semantic properties of the product in specific sensory-semantic systems (see Figure 2). The information stored in these systems may be accessible as long as the product concept is activated. Such conceptual interactions are a result of products being inherently multi-modal objects.

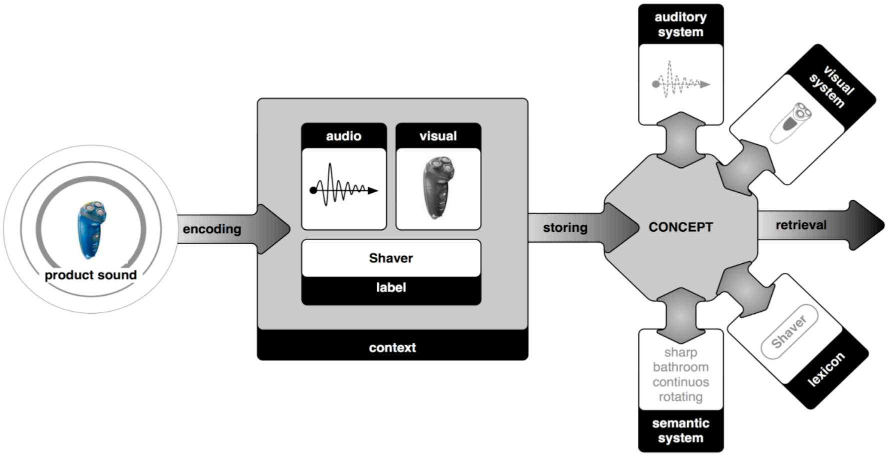


Figure 2. Memory representations of product sounds.

In general, the findings imply a conceptual framework for (product) sound identification that is grounded on empirical findings. The basic construct of the framework stems from theoretical descriptions of object identification. The process of object identification has often been discussed around three main consecutive stages: **perception**, **recognition**, and **identification**. Figure 3 organizes the process of object identification into these three main stages and depicts the mental activities occurring in each stage.

Accordingly, during ***perception*** a featural analysis of the object takes place. At this stage, the structural components of the object are determined via extracting the structural features of the incoming stimuli. Later, these features are integrated to form a structural hierarchy in the object's short-term sensory representation. Finally, a percept of the object is formed (Biederman, 1987; Bregman, 1990; Stevenson & Boakes, 2003). ***Recognition*** is the stage in which the percept is mapped on the long-term memory representations. If there is a match, then recognition occurs. Some theories do not distinguish between recognition and identification (Biederman, 1987; Henderson & Hollingworth, 1999; McAdams, 1993). Yet, they emphasize that recognition / identification entails meaning attribution and a possible access to a lexicon. Other theories consider recognition and identification as two separate phases of meaning attribution (Cleary, 2002; McCauley, Parmalee, Sperber, & Carr, 1980; Peynircioglu, 1990). I consider ***identification*** as the final stage in which meaning attribution occurs. Furthermore, the processes of object identification are prone to top-down (i.e., contextual) influences (Bar, 2004; Henderson & Hollingworth, 1999). Thus, object identification can be biased or guided by mental representations activated by context frames.

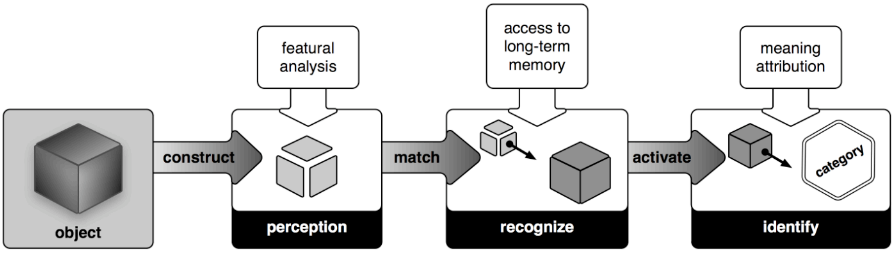


Figure 3. Stages of an object identification process.

These stages of object identification are considered common to the identification of visual, auditory, and olfactory objects (see, e.g., Biederman, 1987; McAdams, 1993; Stevenson & Boakes, 2003). However, procedural differences exist that derive from the nature of the stimuli and from the cognitive system's capacity to encode a specific type of information and make it semantically relevant. Furthermore, each sensory system differs in the way they process and encode the incoming stimuli. For example, the auditory system is quite accurate in capturing the temporal resolution of events (Povel, 1981) and visual system is better at spatial orientation (Kubovy & van Valkenburg, 2001).

New approach to the sound identification process

Historical review of the sound identification processes

Over the years much attention has been paid to the perceptual organization of the sounds in relation to auditory identification. Structural components of the sounds (e.g., pitch, harmonics, rhythm) have been determined in relation to the processing capacity of auditory memory (Deutsch, 1972; Deutsch & Feroe, 1981; Povel, 1981). Sensory processes have been investigated in order to understand how listeners distinguish and mentally represent the auditory qualities of a sound event (e.g., Bregman, 1990; Handel, 1990). Similarly, Gestalt-like rules (e.g., old plus new, proximity, continuity, suddenness) were determined that function in the formation of auditory streams (Bregman, 1978; Bregman, 1990). Listeners' sensory reactions to auditory stimuli have been measured. Consequently, psychacoustical parameters (sharpness, roughness, loudness, and tonalness) have been determined that cause (un)comfortable sensations (Zwicker & Fastl, 1990).

A stage further in the process, meaningful associations of the auditory percept have been considered. For example, semantic correlates of simple tones or natural sounds have been investigated (von Bismarck, 1974; Björk, 1985; Edworthy, Hellier, & Hards, 1995; Solomon, 1958). Thus, its has been revealed that psychoacoustical reactions (e.g., sharp, loud) to sounds in isolation evoke a variety semantic associations in memory (obtrusive, tense, angular, powerful, danger, urgency) other than basic sensory judgments (pleasant-unpleasant). This implies that sounds alone are able to evoke semantically relevant abstract associations which are not necessarily related to source of the sound.

Other studies have investigated whether listeners could identify the cause of an auditory event and the extent to which they can verbalize their auditory experience in the absence of the visual event. It has been demonstrated that listeners are able to describe the material interactions, the action in the event and the size of the object causing the sound (Cabe & Pitternger, 2000; Hermes, 1998; Kunkler-Peck & Turvey, 2000). Although these associations relate to the source of the sounds, they still describe the event causing the sound. Behavioural studies have shown that listeners' main tendency to label an environmental sound is by determining the object causing the sound (e.g., Fabiani, Kazmerski, Cycowicz, & Friedman, 1996). Yost (1991) suggests that the factor in the environmental sound identification is the determination the source of the sound and that happens via accessing the imagery of the source. In addition, identification and labeling nay be two separate processes. In a way, Yost's view is supported by Kubovy and van Valkenburg (2001) that visual domain has an influence in the sound identification. However, this influence on identification and labeling has not been empirically tested for environmental sounds.

Process of sound identification

The literature shows that during identification different stages exists in listener's reactions to a sound, which may be on a perceptual, on an emotional or on a cognitive level. Consequently, sound descriptions may depend on the stages of identification. The studies (Bregman, 1990; McAdams, 1993) that described the processes for sound identification have often focused on perceptual process that the acoustic content of the sound is determined. The actual identification phase is limited with recognition. Meaningful associations are assumed to occur upon recognition and identification is then assumed to be completed. However, not much detail has been provided on how meaning attribution occurs. Recent neuro-psychological studies provide evidence that an almost mandatory conceptual activation occurs upon the perception of a meaningful environmental sound and before labeling (Cummings, Ceponiene, Koyama, Saygin, Townsend, & Dick, 2006; Orgs, Lange, Dombrowski, & Heil, 2006). Conscious identification occurs as a result of the required labeling task. Moreover, top-down (e.g., context) processes have been considered in the existing descriptions of sound identification processes. Context has been used to investigate top-down effects in the sound identification process. However, the provided contexts were restricted to verbal descriptions or different auditory objects (Ballas & Mullins, 1991). No context has been used that provided different types of visual imagery (e.g., picture of the object or the object itself).

Therefore, a revised view on environmental sound identification is necessary in order to understand the meaning attribution to an environmental sound. In this thesis, I propose a sound identification process that focuses more on the cognitive processing of the sound than the perceptual possessing. Because product sounds are ambiguous sounds that are difficult to identify, they may be particularly useful to obtain knowledge on meaning attribution. Meaning attribution may occur at different levels of associations to the source of the sound. That is, listeners may not always directly access the concept of the sound (e.g., shaver) but may access a higher-level categorical association (e.g., bathroom). Therefore, an intermediate stage between recognition and lexical identification should be added. This stage is the *categorical identification* and occurs upon auditory perception and before accessing to a lexicon (c.f. Cummings et al., 2006; Orgs, et al., 2006).

Part A of this thesis has shown that sound identification process does not always operate independently of the information from the visual system. First, it seems that both auditory and visual information are an integral part of the conceptual knowledge of a product. Thus, as soon as the concept of a product is activated (e.g., via a visual context), the relevant auditory, visual, and semantic knowledge will be activated too. Consequently, not only cognitive processes will be influenced by this contagious conceptual activation, also perceptual processes could be hindered or biased by the presence of visual information. Therefore, the positive and negative effects of visual information should be considered on sound identification.

In Figure 4, four stages of the identification process are shown: **perception**, **recognition**, **categorical identification**, and **lexical identification**. Furthermore, in this framework, the **encoding** stage is added as an independent stage that is a pre-requisite to the perceptual and cognitive stages. According to the framework, meaning attribution can occur at any stage; only the specificity of the meaning shifts from sound to sound source throughout the whole identification process. Sound identification becomes completed when the label of the sound source (namely, the product) has been accessed.

Encoding is the stage in which the processed acoustic structure that is stored in the auditory system is encoded and linked to semantic systems (see Chapter 3, Bartlett, 1977; Chiu & Schacter, 1995; Paivio, 1991; Thompson & Paivio, 1994). Sounds that have better structure in their spectral-temporal content (e.g., synthetic musical sounds or machinery sounds with a harmonic and rhythmic structure) have better memory representations, thus are easier to access during recognition. Presence of visual or verbal information at encoding has consequences for further identification processes (Chapter 3). For example, the structure of a sound cannot not be completed if additional visual or verbal information disrupts the encoding of the auditory information (verbal and visual overshadowing effects, see Chapter 3). This disruption negatively influences the recognition stage. However, if a sound is encoded together with a visual and verbal label, then the identification of the sound is easier (dual-coding, see Chapter 3).

During ***Perception***, first, the acoustical content of the incoming sound is extracted to determine the hierarchy in the spectral-temporal structure. Then, these hierarchical components are integrated in order to form an auditory percept. The formation of the sound could be biased by a visual context, if the visual context activates auditory information which is integrated in the context frame (see Chapter 5; c.f. Palmer, 1975; Biederman, 1981). Upon perception, sensory reactions can be elicited and described depending on the acoustical content of the sound (Zwicker & Fastl, 1990). These reactions can occur in the form of emotional (unpleasant, irritating, obtrusive) or psychoacoustical (sharp, loud) descriptions.

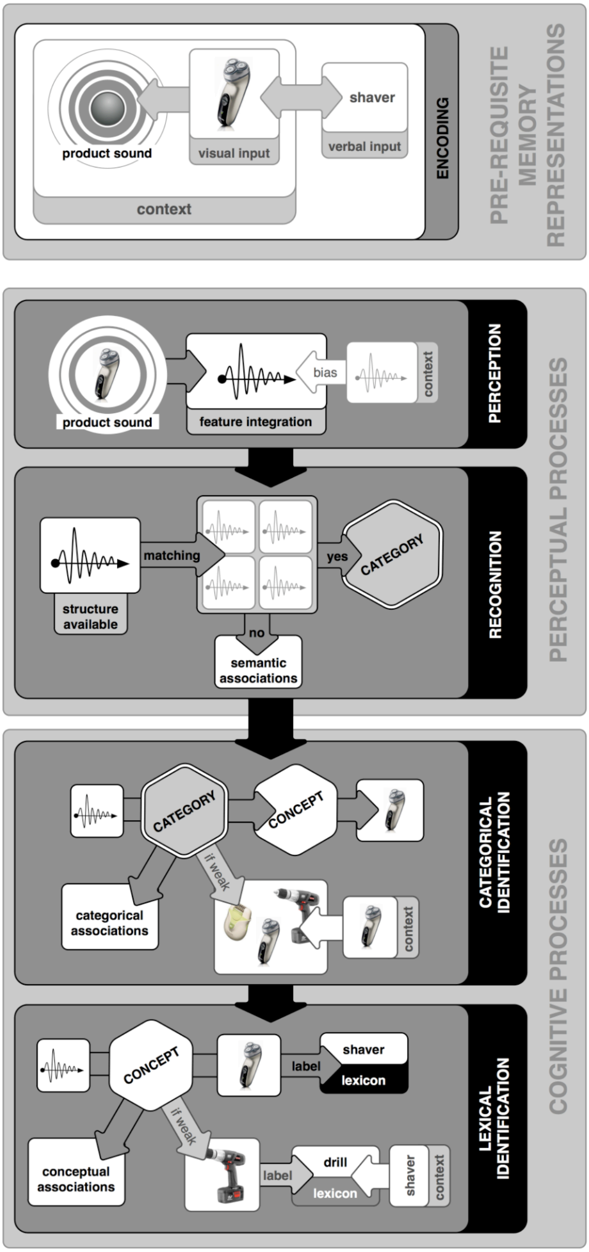


Figure 4. Proposed theoretical framework for the product sound identification process.

Recognition is the stage in which the auditory percept is mapped on to existing auditory representations in long-term memory. If the existing memory representation to which a heard sound is matched is a representative sound of a category, then all other sounds of this category are activated (c.f. Mervis & Rosch, 1981). In other words, if there is a match, sounds within a category are activated Consequently, the sound is recognized, because a similar representation exists in memory. However, the best matching sound representation needs to be determined within this category for the later stages of the identification process. At the recognition stage, certain semantic associations relevant to the acoustical content of a sound (continuous, repetitive, shrill, low-pitched) can occur regardless of whether the sound is recognized (see Chapters 1 & 2; Fabiani et al. 1996; Marcell, Borella, Greene, Kerr, & Rogers 2000). If the sound is not recognized, sound imitations can occur in the form of onomatopoeias (e.g., buzzing).

Categorical identification is an intermediate stage in which for the first time meaningful associations regarding the source of the sound are activated. The category that is activated in the recognition phase provides category-specific mental representations. This may be in the form of super-ordinate level representation of sounds (e.g., locations in which sounds occur, interacting materials and actions that cause the sound) (see Chapters 1 & 2). Categorical activation may not be very salient to a listener, if the recognized sound has strong associations to a concept (e.g., a shaver concept). In that case, the sound will be conceptually identified and then only a lexical identification will be necessary in the next phase. If such a concept cannot be accessed, then this category provides multiple sounds that are similar to the perceived sound (e.g., machinery sound category may provide concepts such as drill, epilator, and shaver) (Chapter 4). This may be the first evidence for ambiguity in sound identification. In the case of ambiguity, a context frame may help to assign a context-relevant concept to the recognized sound (Chapter 5; and also Bar, 2004; Henderson & Hollingworth, 1999). This stage can complete with access to a concept or a category that best represents the recognized sound.

Lexical identification is the last phase of the identification process. At this stage, the lexical representation of a sound is determined by accessing the concept of the sound source. If the sound has strong associations to a concept that represents the sound, then the label of this sound should be active and accessible. If weak associations exist between the sound and the concept, then the presence of a context may improve this (Chapter 5).

Contemplating the proposed framework

Product sounds are, in general, ambiguous sounds when they occur in isolation. Therefore, accessing a lexical representation is often difficult. However, listeners appear to be good at in associating the heard sound to locations or events that may cause it. Accordingly, most of the sound descriptions concern categorical associations. That may be a reason why such an additional stage is included in the identification process. However, sound identification may occur very fast and accurate, if the sound has clear and easy-to-access memory representations. Such fast activation of the product concept may overshadow the categorical activation. However, once the sound is identified, its concept will activate a conceptual network that regards both source and sound related (higher-level) associations.

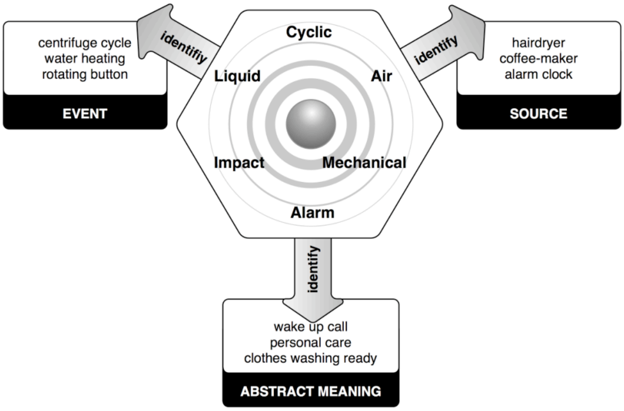


Figure 5. Main product sound identification types.

Determining the end-result of an identification process may be challenging because the events and actions causing a product sound are not always visible to a listener. One may identify a sound as a kitchen extractor, because the sound is conceptually and experientially associated to the product (but not because the mechanisms causing the sound event are visible). Consequently, a sound may be considered as a property to a product. Therefore, providing a sound label will evoke an activation of a product concept before the label can be retrieved. However, this thesis has provided evidence that not all product sounds are identified by the source (i.e., product) causing the sound. Depending on the type of sound, different sound identification types occur (see Figure 5). For alarm sounds meaning of the alarm needs to be extracted. For example, an alarm sound could be best described by “food is ready” if it is a microwave oven bell sound, or as “wake up call” if it is an alarm clock sound. As for the sounds caused by events that are visible and that require product-user interaction (impact and liquids sounds), the representations in visual and motor systems could be complementary to the auditory experience in encoding the meaning of the sound. Thus, such sounds could be labeled as water boiling or door closing.

For the proposed framework, sound identification is assumed to be completed when a sound source is accessed. That is because, labels provided for product sounds reflect mostly the name of the product. Future studies should incorporate other sound identification types (i.e., event and abstract meaning), because different perceptual and cognitive mechanisms may be operating for these identification types. For example, processing of the alarm sound may be more on a perceptual level and meaning attribution may stop at perceptual or categorical identification stage. The use of source information *may* be redundant for such sounds, because they may have associations to abstract meanings. For event related sounds, visual input may be essential to be able to perfectly identify them. Such sounds were shown to be rather ambiguous sounds in isolation. However, with visual information that associates the sound to an interaction event, the identification seems to improve (see Chapter 5).

Emotional responses were mentioned to occur during perceptual processes in the proposed framework. However, we found evidence (Chapter 1, Experiment 5) that there may be cognitive reasons for the emotional responses. For example, in isolation, the sound of an epilator may be perceived rough and therefore a little discomforting. However, knowing the function of an epilator may totally change our perception of the sound because of the previous ‘painful’ experiences with the product. Therefore, a distinction should be made between psychacoustical responses resulting from the sound’s auditory qualities and cognitive-driven emotional responses resulting from the concept of the product.

References

Ballas, J. A., & Mullins, T. (1991). Effects of context on the identification of environmental sounds. *Human Performance, 4*(3), 1991 - 1219.

Bar, M. (2004). Visual objects in context. *Nature Reviews: Neuroscience, 5*(8), 617-629.

Bartlett, J. C. (1977). Remembering environmental sounds - Role of verbalization at input. *Memory & Cognition, 5*(4), 404-414.

Biederman, I. (1981). On the semantics of a glance at a scene. In M. Kubovy & J. R. Pomerantz (Eds.), *Perceptual Organization* (pp. 213 - 263). New York: Lawrence Erlbaum.

Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review, 94*(2), 115-147.

Bismarck, G. (1974). Timbre of steady sounds: A factorial Investigation of its verbal attributes. *Acustica, 30*, 146 - 159.

Björk, E. A. (1985). The perceived quality of natural sounds. *Acustica, 57*(3), 185-188.

Bregman, A. S. (1990). *Auditory Scene Analysis: The Perceptual Organization of Sound* Massachusetts: MIT Press.

Bregman, A. S. (1978). The formation of auditory streams in J. Requin (Ed.) Attention and Performance VII. Hillsdale, NJ: Lawrence Erlbaum.

Cabe, P., & Pittenger, J. B. (2000). Human sensitivity to acoustic information from vessel filling. *Journal of Experimental Psychology: Human Perception and Performance, 26*(1), 313 - 324.

Chiu, C. Y. P., & Schacter, D. L. (1995). Auditory priming for nonverbal information: Implicit and explicit memory for environmental sounds. *Consciousness and Cognition, 4*(4), 440-458.

Cleary, A. M. (2002). Recognition with and without identification: Dissociative effects of meaningful encoding. *Memory & Cognition, 30*(5), 758-767.

Cummings, A., Ceponiene, R., Koyama, A., Saygin, A. P., Townsend, J., & Dick, F. (2006). Auditory semantic networks for words and natural sounds. *Brain Research, 1115*(1), 92-107.

Deutsch, D. (1972). Mapping of interactions in the pitch memory store. *Science, 175*(25), 1020-1022.

Deutsch, D., & Feroe, J. (1981). The Internal Representation of Pitch Sequences in Tonal Music. *Psychological Review, 88*(6), 503-522.

Edworthy, J., Hellier, E., & Hards, R. (1995). The Semantic Associations of Acoustic Parameters Commonly Used in the Design of Auditory Information and Warming Signals. *Ergonomics, 38*(11), 2341-2361.

Fabiani, M., Kazmerski, V. A., Cycowicz, Y. M., & Friedman, D. (1996). Naming norms for brief environmental sounds: effects of age and dementia. *Psychophysiology, 33*(4), 462-475.

Handel, S. (1989). *Listening*. Massachusetts: MIT Press.

Henderson, J. M., & Hollingworth, A. (1999). High-level scene perception. *Annual Review of Psychology, 50*, 243-271.

Kubovy, M., & Van Valkenburg, D. (2001). Auditory and visual objects. *Cognition, 80*(1-2), 97-126.

Kunkler-Peck, A. J., & Turvey, M. T. (2000). Hearing Shape. *Journal of Experimental Psychology: Human Perception and Performance, 26*(1), 279 - 294.

Marcell, M. E., Borella, D., Greene, M., Kerr, E., & Rogers, S. (2000). Confrontation naming of environmental sounds. *Journal of Clinical and Experimental Neuropsychology, 22*(6), 830-864.

McCauley, C., Parmelee, C. M., Sperber, R. D., & Carr, T. H. (1980). Early extraction of meaning from pictures and its relation to conscious identification. *Journal of Experimental Psychology: Human Perception and Performance, 6*(2), 265-276.

McAdams, S. (1993). Recognition of sound sources and events. In A. S. Adams & E. Bigand (Eds.), *Thinking in sound: The cognitive psychology of human audition* (pp. 146 - 198). New York: Oxford University Press.

McCauley, C., Parmelee, C. M., Sperber, R. D., & Carr, T. H. (1980). Early extraction of meaning from pictures and its relation to conscious identification. *Journal of Experimental Psychology: Human Perception and Performance, 6*(2), 265-276.

Orgs, G., Lange, K., Dombrowski, J. H., & Heil, M. (2006). Conceptual priming for environmental sounds and words: an ERP study. *Brain and Cognition, 62*(3), 267-272.

Paivio, A. (1991). Dual coding theory - Retrospect and current status. *Canadian Journal of Psychology-Revue Canadienne De Psychologie, 45*(3), 255-287.

Palmer, S. E. (1975). The effects of contextual scenes on the identification of objects. *Memory & Cognition, 3*(5), 519-526.

Peynircioglu, Z. F. (1990). A feeling-of-recognition without identification. *Journal of Memory & Language*, 29, 493-500.

Povel, D. J. (1981). Internal representation of simple temporal patterns. *Journal of Experimental Psychology:*

Human Perception and Performance, 7(1), 3-18.

Thompson, V. A., & Paivio, A. (1994). Memory for pictures and sounds: independence of auditory and visual codes. *Canadian Journal of Experimental Psychology, 48*(3), 380-398.

Yost, W. A. (1991). Auditory image perception and analysis - The basis for hearing. *Hearing Research, 56*(1-2), 8-18.

Zwicker, E., & Fastl, H. (1990). *Psychoacoustics: Facts and Models*. Berlin: Springer.

PART B

⁂

APPLICATION

This chapter is based on the paper:

Özcan, E., & van Egmond, R. (2004). ***Pictograms for sound design: A language for the communication of product sounds.*** **Paper presented at the 4th International Conference on Design and Emotion, Ankara, Turkey.**

Abstract

Sound designers often encounter problems in communicating a newly developed sound concept to other designers, researchers, engineers, and marketeers. To support the communication of product sounds, a pictographic language has been designed to describe product sound events in such a way that it relates the physical properties of sounds to its perceptual properties. This new pictographic language will be presented and it will also be discussed how it facilitates the communication of a product sound.

CHAPTER 6

⁂

PICTOGRAMS FOR SOUND DESIGN: A LANGUAGE FOR THE COMMUNICATION OF PRODUCT SOUNDS

Visual information design

Visual information design is a very expansive field, which comprises interface design, music notation, warning and instruction design, sign language for hearing-impaired, etc. It provides fast and accurate communication in user interfaces and employs graphical representation systems such as icons, pictograms, and symbols. The amount of information and similarity decreases from icons to symbols compared to the original object/event. On the one hand, icons and pictograms represent events and objects within a context (e.g., Olympic games pictograms). On the other hand, symbols are abstract and arbitrary representations that require learning (e.g., the letters of the alphabet, the traffic sign ‘no parking’); and they might not refer to any existing objects or events. A visual language makes use of these graphical representation systems to enhance the communication. A spoken or written language would need much more elements (i.e., words) to indicate a certain object or event, whereas in a pictographic language, only one or two elements are needed, the information is more condensed or it is represented by a set of graphical features that are very typical. In addition, it enables communication between people with a different linguistic background. For example, in the airports, a right-arrow sign and a suitcase pictogram together construct a visual sentence, which means that the baggage service is located on the right hand side.

The term ‘icon’ has become to be associated with any functional images used in GUI (Graphical User Interface) designs after the emergence of desktop metaphors (e.g., Macintosh Finder). Icons represent programs, folders, functions, menus, etc. Interface designers have explored the use of sound in GUIs, as well. Gaver (1989) introduced the term ‘auditory icons’ for his SonicFinder design (an auditory interface developed for Apple Computer, Inc). His auditory icons represent the physical sounding objects (e.g., a trash can sounds like something fell into a metal trash can) and their function is similar to visual icons. Blattner, Sumikawa, and Greenberg (1989) used the term called ‘earcons’ to indicate abstract sounds or rhythmic musical patterns. Mynatt (1994) developed a design methodology for auditory icons; identifiability, conceptual mapping, physical parameters, and user preferences are the factors in this methodology. Edworthy and Adams (1996) pointed out that in the design of warning symbols, legibility, conspicuity, discriminability, and urgency mapping is required in order for people to comprehend and to learn the symbols. Holmes (2000/2001) discerned various conventions to design a set of pictograms (e.g., the overall shape, the style of drawings, the subject matter, the context, and the color).

Until now, the work in auditory icons has dealt with the sonification of the graphical objects or events in GUIs. Whereas in this study, pictograms, which belong to a certain sound aspect, have been designed to develop a pictographic language that visualizes sounds. The possibility of using of pictograms in a visual language will be explored in this paper. The sounds of products and the sound producing parts of products will be taken as a starting point for the design of pictograms. Therefore, the perceptual and physical qualities of sound and image should be investigated.

Product sound quality

Product sound design incorporates the sound in the design of a product to increase the product acceptability amongst users. Pursuing the new developments in engineering, technology, production techniques, chemistry of the materials, and marketing strategies, a product sound designer tries to estimate the users’ expectation of the product -at the design stage- and to adapt the sound quality to a desired one. In addition, because of the new insights in human perception and emotional experience of sound, producers have come to realize that product sounds have high-level affective influences on user experiences of a product. For example, adjusting the sound pressure level in car interiors may influence users’ emotional experiences upon the car. The new sound might indicate a pleasant ride, an expensive car, or a tolerable noise level.

Traditionally, product sounds have been treated as noise; and engineers have only been concerned with the reduction of the sound pressure level emitted by a product (noisy product) to a level that it does not annoy users. Recently, the complexity of the requirements to design a product as well as a product sound has increased due to the producers' demands on cost effectivity, time efficiency, environmental friendliness, up-to-date design; and due to the new insights in human perception, cognitive processes, and emotional experience. Furthermore, recent available technology allows designers to seek and discover the cause of an undesirable sound. Subsequently, they replace the problematic part of the product with a suitable one.

In the recent definition of product sound quality, Blauert and Jekosch (1997) argue that product sound quality is a result of judgments on auditory characteristics of a product sound performed by a user. The mood of a user may influence the perceived product quality, as well as the perceived product sound quality may influence the actual emotional state of the user depending upon the cognitive judgments.

The recent definition of product sound quality contradicts the traditional view in the sense that the sound caused by a product cannot always be treated as noise. Moreover, it may signify functionality, feedback, luxury, comfort, ease, attention, and brand value depending on user expectations of a product. In a user study, the results showed that the truck drivers prefer to receive feedback sound from the engine while they are driving instead of a sound attenuated truck interior design (Talamo, 1982). Moreover, brands like Harley Davidson or Grolsch especially use sound as a brand value, which may exemplify the impact of product sounds on users.

Sound quality in physical terms

Sound is the result of fluctuations in air pressure. It is expressed in decibels (sound intensity), amplitude (sound pressure), frequency (the variation rate of air pressure), etc. Sound quality is often described in terms of psychoacoustical measures. Loudness, roughness, noisiness, sharpness, brightness, and pitch are some of the commonly used psychoacoustical attributes that determine the auditory sensory pleasantness (Zwicker and Fastl, 1990). To measure these psychoacoustical attributes, the physical parameters of product sounds need to be known. By doing so, a product sound designer can determine what range of sounds elicits the auditory sensory pleasantness (or unpleasantness).

Sound is a time-dependent percept. Therefore, the physical parameters of a sound are calculated over time and depicted on a timeline. Waveforms reflect the pressure variations as a function of time; and sonograms show the frequency distribution in time. However, these representations do not disclose the functions and the usage process of products. In addition, it is almost impossible to derive perceptual or experiential aspects of a sound from the shape of its waveform. Although, these visual representations facilitate the communication of product sound to a certain extent in physical terms, such representations still require the perceptually related information flow.

Perceived sound quality

Not much is known about how people perceive product sounds. However, several researchers investigated the timbral properties of sound. Solomon (1958) found relevant descriptive adjectives that could be used to characterize passive sonar sounds. Von Bismarck (1974) found perceptual dimensions underlying verbal attributes that described the timbre of steady state sounds, of which sharpness was the most important. Björk (1985) investigated the emotional dimensions associated with auditory sensation. The results revealed that roughness, sharpness-pitch, and loudness sensations were correlated with evaluation, activity, and potency dimensions respectively. Kendall and Carterette (1993) used the verbal attributes of Von Bismarck's adjectives to describe the timbre of musical instruments, and concluded that those adjectives were not appropriate for the purpose. It has also been argued that psychoacoustical terms 'sharpness' and 'roughness' were associated with 'pleasantness' of the sound (Zwicker and Fastl, 1990).

Other researchers have investigated the identification of sound and its sources. Bregman (1990) proposed two important concepts that underlie the mental process of how people come to perceive events: auditory stream segregation, and/or auditory stream integration. McAdams (1993) discussed in detail the auditory recognition process and suggested that auditory recognition and auditory identification are two separate but consecutive stages. With respect to the cognitive aspects of sounds, Bartlett (1977) investigated the role of verbalization in memory for environmental sounds and showed that people could recognize the labels and sources of environmental sounds. In another study, Ballas (1993) showed that identification time, occurrence frequency, and cognitive aspects of everyday sounds play important roles in identification of sources of everyday sounds. A recent study by Kunkler-Peck and Turvey (2000) indicated that one could hear the shape, size, and material of thin plates.

Gaver's (1993a) research on the categorization of sounds is of interest for the present study. He discerned three categories for sound producing events: sounds of vibrating objects, liquid sounds, and aerodynamic sounds. Still, this categorization may not be sufficient to define the domain of product sounds. Gaver (1993b), also, described the physics of sound-producing events to provide an initial orientation towards the relevant attributes of sound-producing events. He concluded that different type of impacts elicits different frequency range of sounds depending on the material, stiffness, and medium of the impacting objects.

The interdisciplinary communication problem of product sounds

In the course of the product design process, experts from several areas provide knowledge to contribute to the overall design project. To improve the design quality, a continuous information flow can be observed among these experts. An industrial engineer is responsible for the feasibility of the whole project, whereas a mechanical engineer may deal with the manufacturing techniques regarding the material choice. An industrial design engineer decides how the product should function and what kind of mechanical/electrical part would result better in the product functionality. A marketeer defines the marketing strategies for the product acceptability amongst the target group. An advertiser communicates the values of the product to the potential users.

In this team, a sound designer participates in the design process of a product when a new sound concept needs to be introduced to the overall design concept of the product. The task of the product sound designer may vary from solving the roughness problem in the sound of a domestic appliance to translating the total design concept of a product into sound design. By doing so, the sound designer needs to communicate with engineers about, for example, the engine materials in order to evoke efficiency, expensiveness, comfort, etc. In the following example, a description of the sound design problems of a product is presented (Fog and Pedersen, 1999).

> *“A car-maker wanted a silent power-steering with a faint quality sound. The sub-supplier asked for an analysis of the sound from the existing power-steering systems, and specifications of the desired “sound” and suggestions to design changes. As a result of this project, new owners of that make can pride them selves on the quiet and harmonious sound of their power-steering.”*

The multidisciplinary nature of the product design process requires precise, effective, and dynamic information flow; therefore, no vague, misspelled, arbitrary communication types can be afforded. However, it is difficult to describe the properties of sounds, either verbally or graphically, because a common verbal language (i.e., lexicon) is missing. For example, descriptive words like ‘rough’, ‘soft’, or ‘round’ do not immediately relate to a specific property in sound. Many products exist causing ‘rough’ sound such as shavers, epilators, etc. Moreover, a semantic problem exists in sound descriptions. Rough and soft are tactile sensory percepts, and round is a visual one. It was also shown (Martens and Giragama, 2002) that the words used to describe guitar timbres in Japanese and Sinhalese languages (with the same English meanings, e.g., pleasant, cheerful, sharp) were related to different acoustical dimensions in those languages. This may indicate a language problem. Yet, a sound designer needs to discuss some attributes of the product sound with the engineers, marketeers, or other designers who are involved in the design process. Supposedly, this multidisciplinary team has insight in the overall design concept of the product; however when discussing the sound-specific topics, they might not have sufficient background in auditory event perception and familiarity with sound-related technical terms (e.g. bark scales, modulation frequency, sones, phons, onset/offset envelopes, etc.).

Proposed solution:

A pictographic language for the communication of product sounds

Why should the use of pictograms solve the above-mentioned communication problem? It is obvious that no verbal or graphical language can replace the percept of a sound. Hearing an actual sound evokes a different perceptual experience or associated meaning than having to imagine the sound on the basis of a verbal description. Therefore, the proposed pictographic language is not supposed to substitute existing percepts; instead, it is a design proposal to improve the existing communication for product sounds. A combination of sound and a pictographic language may communicate in a faster and more accurate way.

Pictograms for product sound design should represent the source that generated the sound. To be able to design such pictograms, insight in physical and perceptual product sound characteristics is required. In the physical space, product sounds originate from the moving (sliding, rotating, rinsing, etc.) and contacting (hitting, scraping, dropping) parts of the products, the engines, fans, buttons, etc. The frequencies radiated by a product depend on the material stiffness and the medium of the vibrating objects. The amount of force and the resonance properties applied to the moving or static product parts determine the sound pressure level (amplitude). In the perceptual space, product sounds are described in terms of relevant perceptual attributes. In auditory perception, it has been shown that people were able to perceive the material, size, and shape of sounding objects and describe the perceptual qualities of sounds (Hermes, 1998; Kunkler-Peck and Turvey, 2000; Klatzky, Pai, and Krotkov, 2000). Furthermore, product sounds elicit certain emotions that influence the total product perception (Västfjäll, 2002).

Pictograms and categorization

In a pictogram domain, the pictograms represent the sounds. As a first step, the domain of product sounds should be defined. Decomposing a product into its functions and its parts shows the type of sounds and sound sources are involved. For example, a vacuum cleaner has a plug for the electricity, power button to turn on/off, and speed adjustment button. These parts need interaction with the user and the sound comes out during or after the interaction. The engine and fan sound of the vacuum cleaner is generated at the stage of sucking and blowing air. The sucking/blowing sound changes its property depending on the speed of the fan and engine. These types of engine sounds are intrinsic sounds; they occur when the product is running. After decomposing the products into its functions and its parts, the underlying perceptual product sound categories should be defined. This requires the listeners' perceptual and cognitive judgments upon the product sounds. Therefore, in a previous study, we investigated the underlying dimensions of perceptually relevant product sound categories (Chapter 1). Do users categorize by sound source, by product type, by sounds' physical or psychophysical attributes, by affective responses? These aspects will determine the type of pictograms (including icons, pictograms, and/or symbols) used in the pictographic language.

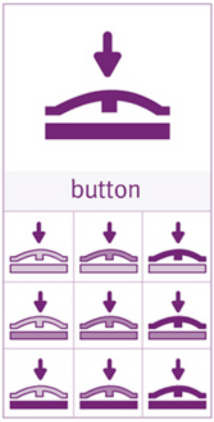


Figure 1, Relationship among pictograms in a group.

In a category, the extent of similarity among the members differs. Therefore, the difference in similarity among members should be reflected in the design of the members of the pictogram groups. Each group covers certain product sounds on different dimensions (i.e., the material interaction type may categorize button sounds; whereas, the amount of impact absorbed by the material and the material stiffness may categorize impact sounds). One member of a category exists that is more similar to all other category members. This member may represent the whole category (Mervis and Rosch, 1981). Consequently, the most representative pictograms should be designed in such a way that they represent the groups. Mervis and Rosch observed a hierarchical organization among categories.Similarly, the pictogram categories are organized at a basic level and a sub-ordinate level. The possible organization of pictograms is shown in Figure 1.

The scenario

Imagine that the following problems with sound have been encountered in the design process of a vacuum cleaner, and the sound designer has to explain this problem to the project manager.

> *"The usability tests revealed that the vacuum cleaner has an unexpected loud sound just 2-3 seconds after it starts running. Most of the users had the impression that the machine got broken at that time, hearing the sound was annoying, and the sound was uncontrolled. The sound designer doesn't know which part causes the loud sound; however, he has to fix it because the vacuum cleaner should sound expensive, accurate, and pleasant."*

As a first step, the sound designer records the sound of the vacuum cleaner to analyze the physical properties of the sound. The recorded sound (Figure 2) presents the relative amplitude of the vacuum cleaner sound as a function of time. This figure may reflect several events. In the figure, one short and one long event can be observed. The first event seems to have happened very abrupt and decayed in less than 100 milliseconds. The second event seems to have started fading-in and finished fading-out slowly. In addition, another peak, similar to the first event, has been observed in the second event near the end of the recording. Approximately at the 2nd second, the increased amplitude can bee seen in the figure, and this indicates the users' complaints about the vacuum cleaner sound. However, this figure does not convey the auditory qualities of the sound. It is hard to imagine how this waveform would sound like. One cannot distinguish the sound of a vacuum cleaner from the sound of, for example, a microwave oven by just looking at the waveform of a sound recording.

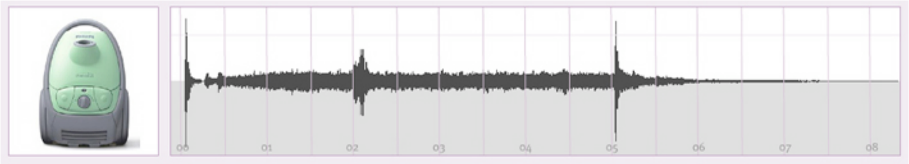


Figure 2. Amplitude representation of the vacuum cleaner sound.

In the second step, it is necessary to determine the perceptually relevant components of a product sound. The possible sound sources should be determined to find the part that makes the undesired sound. Therefore, the vacuum cleaner is decomposed into parts that generate sound. Three main sound sources exist in the function cycle: an on/off button, an engine, and a fan. The sound designer records the sounding parts separately to find out the problem. The user turns on the machine by pressing the on/off button. The engine and the fan start running, subsequently. The sound of an on/off button is followed by the engine and fan sounds. The user turns the machine off by pressing the button again. The engine and the fan stop running; but the fan sound takes longer to stop. After determining the sound sources and the processes, the sound designer selects the proper pictograms for the sounding parts. In Figure 3, the sound designer finally attaches the sound pictograms to the physical properties of sound (i.e., amplitude, in this case). The to-be-analyzed sound property can, of course, be replaced by other physical (e.g., sonogram) or psychophysical attributes (e.g., brightness, roughness, pitch etc.) properties of a sound.

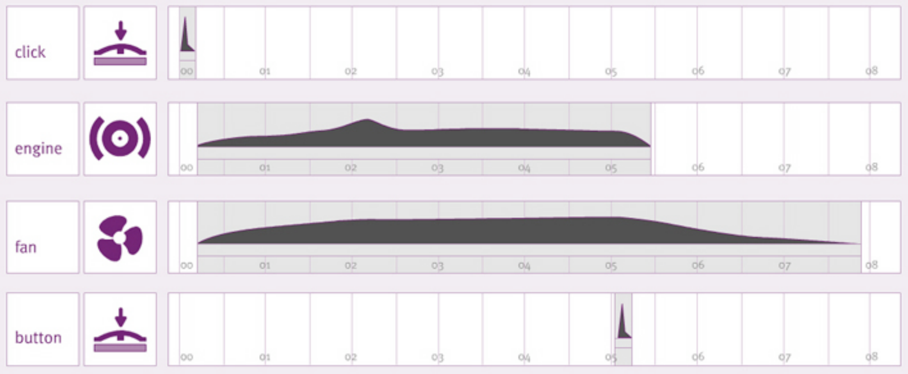


Figure 3. Pictographic representation of the amplitude of the vacuum cleaner sound.

Once the whole product sound is represented visually, the sound designer refers to the problem in a timeline. It was said that the loud sound started 2-3 seconds after the vacuum cleaner starts running. It seems that at that time, only the fan and the engine generates sound. However, the engine sound has a drastic change in the amplitude envelope at around 2nd second. Once the whole sound events and the problem is visualized by pictograms, the visualization of the sounds events can be printed and be used as a guide to show the problem. In design meetings, such a document may also serve as a reference tool for the designers, where there are only paper documents are available.

The design of the pictograms

The proposed pictographic language is still under development. Thus, some ideas may seem very vague. However, some effort has been put in designing some of the pictogram groups to see whether the proposed language would be applicable. The dimensions, which underlie the design criteria of a pictogram group, differ from one group to another because of the multidimensional nature of product sound quality. In Figure 1, the button sounds were visually represented. The button sounds are the cause of two interacting materials, and the stiffness of these materials determines the click sound. The metal-to-metal clicks generate a higher frequency sound than the plastic-to-plastic clicks. In the Figure 1, from top to bottom, stiffness of the lower interacting material increases, whereas, from left to right, stiffness of the upper interacting material increases. In Figure 4, for the same pictograms, the tone of the color indicates the stiffness of the material. The darker colors indicate hard materials (e.g., metal, glass), the lighter colors indicate softer materials (e.g., rubber, plastic). These pictograms may represent, different kinds of button, and click sounds.

Figure 4. Various button pictograms.

A categorization study might result in perceptual product sound groups and the hierarchy among them. In Figure 4, the possible subordinate level product sounds have been displayed, whereas, the possible basic level product sounds are shown in Figure 5.

Figure 5. Basic level representative pictograms.

Other application areas

Industrial designers and engineers often use CAD-systems to model their designs in 3D environments. 3D modeling is necessary for the first impressions of the design project. Generally, 3D modeled product designs expose an almost real look-and-feel of the product. Often, animations are used to show how the product works in virtual environment. However, these animations are not supported by product sounds. If the sound percept is integrated in the presented work, the 3D model would be more realistic and convincing. In this way, by using the pictogram domain and corresponding sounds, industrial designers could model the soundscape of the product, as well. For this, creating a sound library for pictogram groups would be necessary.

Another form of application would be at the sound concept development stage. Designers could use the same sound-pictogram library. By arranging the pictograms (with sounds) on a timeline with relevant psychoacoustical measures, they could design the desired soundscape of a product as a product sound concept proposal.

Discussion: Pictograms for emotions

The above-proposed visual language excludes the emotional impact of sounds. However, the definition of product sound quality suggests that the product sounds elicit emotional responses. Therefore, the experienced emotion evoked by the product sound becomes one of the descriptive attributes of sound and needs to be considered in the design process of a product sound.

Suggested is that pictograms for emotions may represent the emotion domain for product sounds. Desmet (2003) designed a product emotion measurement tool by employing caricaturistic animations of a puppet, which express emotional experience of a product. Van Egmond (2004) used this tool to measure elicited emotions on alarm sounds. It seems possible to use abstract representations (graphical) to express emotions. Furthermore, analyzing the product sound beforehand may enable the sound designer to predict the users’ possible emotional responses upon the experienced sound.

Acknowledgements

We would like to acknowledge Piet Westendorp’s help for providing us with knowledge on visual information design and thank him for his comments on the design of the pictographic language.

References

Ballas, J. A. (1993). Common factors in the identification of an assortment of brief everyday sounds. *Journal of Experimental Psychology: Human Perception and Performance, 19*(2), 250-267.

Bartlett, J. C. (1977). Remembering environmental sounds: The role of verbalization at input. *Memory & Cognition, 5*(4), 404-414.

von Bismarck, G. (1974). Timbre of steady sounds: A factorial investigation of its verbal attributes. *Acustica*, 30, 146-159.

Björk, E. A. (1985). The perceived quality of natural sounds. *Acustica, 57*, 185-188.

Blattner, M. M., Sumikawa, D. A., & Greenberg, R. M. (1989). Earcons and icons: Their structure and common design principles. *Human Computer Interaction, 4*, 11-44.

Blauert, J. & Jekosch, U. (1999). Sound quality evaluation: A multi layered problem. *Acustica, 83*, 747-753.

Bregman, A. S. (1990). *Auditory scene analysis: The perceptual organization of sound.* Massachusetts: MIT Press.

Desmet, P. M. A. (2002). Designing emotions. Doctoral Dissertation, Delft University of Technology, Delft, The Netherlands.

Edworthy, J., & Adams, A. S. (1996). Symbols. In *Warning design: A research prospective*, (pp. 75-100). London: Taylor & Francis.

Fog, L., & Pedersen, T. (1999). Tools for product optimization. *Human Centered Processes 10th Mini EURO Conference*, Brest, France.

Gaver, W. W. (1989). The SonicFinder: An interface that uses auditory icons. *Human Computer Interaction, 4*, 67-94.

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

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

Hermes, D. J. (1998). Auditory material perception. *IPO Annual Progress Report, 33*, 95-102.

Holmes, N. (2000). Pictograms: A view from the drawing board, or what I have learned from Otto Neurath and Gerd Arntz (and Jazz). *Information Design Journal, 10*(2), 133-143.

Kendall, R. A. & Carterette, E. C. (1993). Verbal attributes of simultaneous wind instrument timbres: I. von Bismarck's Adjectives. *Music Perception, 10*(4), 445-468.

Klatzky, R. L., Pai, D. K., & Krotkov, E. (2000). Perception of materials from contact sounds. *Presence, 9*(4), 399-410.

Kunkler-Peck, A. J., & Turvey, M. T. (2000). Hearing shape. *Journal of Experimental Psychology: Human Perception and Performance, 26*(1), 279-294.

Martens, W. L. & Giragama, C. N. W. (2002). Relating multilingual semantic scales to a common timbre space. *113th Convention of the Audio Engineering Society*, Los Angeles, USA.

McAdams, S. (1993). Recognition of sound sources and events. In S. McAdams and E. Bigand (Eds.) *Thinking in sound: The cognitive psychology of human audition*, (pp. 146-198). New York: Oxford University Press.

Mervis, C. B., & Rosch, L. (1981). Categorization of natural objects. *Annual Review of Psychology, 32,* 89-115.

Mynatt, E. D. (1994). Designing with auditory icons: How well do we identify auditory cues? *Chi '94 Conference Companion, 269-270*. New York: ACM.

Özcan, E., Van Egmond, R., & Jacobs, J. (submitted). Categorization and identification of product sounds.

Solomon, L. N. (1958). Semantic approach to the perception of complex sounds. *The Journal of the Psychoacoustical Society of America, 30*(5), 421-425.

Talamo, J. D. C. (1982). The perception of machinery indicator sounds. *Ergonomics, 25*, 41-51.

Van Egmond, R. (2004). Emotional experience of frequency modulated sounds: implications for the design of alarm sounds. In D. de Waard, K. A. Brookhuis & C. M. Weikert (Eds.), *Human factors in design* (pp. 345-356). Maastricht: Shaker Publishing.

Västfjäll, D., Gulbol, M. A., Kleiner, M., & Gärling, T. (2002). Affective evaluations of and reactions to exterior and interior vehicle auditory quality. *Journal of Sound and Vibration, 255*(3), 501-518.

Zwicker, E., & Fastl, H. (1990). *Psychoacoustics: Facts and models*. Berlin: Springer.

This chapter is based on the paper:

Özcan, E., & van Egmond, R. (2006). Product sound design and application: An overview. Proceedings of the International Conference on Design and Emotion, Sweden, 5, 1-19.

Abstract

Designing sounds for products is a relatively new concept that still lacks a systematic procedure. Formerly, a pleasing sound would be realized by, e.g., decreasing the sound pressure level. The sound of a vacuum cleaner would be changed if it sounded unpleasantly loud. Nowadays it is acknowledged that product sound design should be integrated in the main design process to enhance the user experience both on ergonomic and hedonic levels. For example, the sound producing parts of an expensive car (i.e., engine, doors, gearbox, cabin auditory warnings, etc.) are designed to reflect the main requirements for the car design such as reliability, comfort, safety, luxury, and consequently a pleasant drive. However, an integration of the sound design into the design process of a product requires specific tools and methods. This paper addresses this issue and presents two methods and a tool by which designers can model and demonstrate their conceptual ideas for the sound of the product under development.

CHAPTER 7

⁂

PRODUCT SOUND DESIGN AND APPLICATION: AN OVERVIEW

Need for Product Sound Design

Alberto Alessi stated in a recent interview on Euronews (Le Mag, 21 April 2006) that when buying products users do not seek functionality anymore—nowadays functionality is taken for granted in any mass-produced product—they rather seek experience, fun, pleasure, and comfort while using a product. This view corresponds to Desmet's (2002) basic model of product emotions which suggests that users may have various concerns with respect to the product use and these concerns may evoke emotional responses. One of these concerns might be related to the sound property of the product which has indeed the potential to influence users' behavior and appreciation of products.

Sounds' influence on users may directly stem from the spectral-temporal composition of the sound itself (Västfjäll et al., 2002) or it may also stem from concurrent cognitive judgments both on the sound and its relation to its source (i.e., the product) and from the meaning derived from this relationship (Özcan & van Egmond, 2005). An example of the spectral-temporal influence would be the sound of a vacuum cleaner. This sound may evoke negative emotional responses on a user due to its high-pitch and loudness level. An example of the cognitive influence would be the sound of a coffee maker. Despite its noisiness and sharpness in the spectral composition, this sound may evoke positive emotional responses on a user due to anticipation of a relaxed time while drinking coffee. Some studies have already shown the emotional effects of product sounds on users (Västfjäll, 2002; Västfjäll et al., 2002; Västfjäll et al., 2003). For domestic appliance sounds, emotional responses often lay on the negative side of the emotion spectrum (Özcan & van Egmond, 2005) which already creates new avenues for product sound related research and design.

In marketing, a product’s visual property is the most important aspect for the first-moment-of-truth (i.e., shelf-presence of products) to convince the potential user to choose between the two similar products from different brands. However, users’ satisfaction with products depends highly on the properties of the product that are available only in the second-moment-of-truth (i.e., product use). Sound as a property naturally emerges only when the product is working. In the course of the product usage, users experience whether or not the produced sound is fitting the total design and the function of the product (Blauert and Jekosch, 1997). So, a sound’s inappropriateness to the product may cause user dissatisfaction. Therefore, not to have any unconsidered effects, product sound should be designed to fit the product values.

Product sounds

We define product sounds as the sounds that are emitted by products as a result of their functionality. For example, a vacuum cleaner makes sound because of the running engine, rotating fan, and air-flow in the tubes. As the name 'product' embodies various artifacts that are mass-produced and entered in the consumer market, product types may vary in the domains of packaging, personal care products, food, and domestic appliances. Relevant to us is the domain of domestic appliances. In this domain, we have defined six perceptually relevant product sound categories (see Chapter 1). These categories are air, alarm, cyclic, impact, liquid, and mechanical sounds. People react differently to the sounds in these categories and their concerns vary from one category to another. For alarm sounds the derived meaning (e.g., ‘food is ready’ for the alarm of the microwave oven) is important to a user, for mechanical sounds the type of source becomes important.

Product design

The main lack in product sound design is a systematic methodology that organizes the sound related design activity. Sound designers also lack tools that assist them creating new ideas and communicating about the sound of the product in development. However, methods from product design can be applied to product sound design. Thus, in order to fill these gaps, we analyzed existing methods for product design in terms of product development processes and design communication. Figure 1 organizes the relationship between the product development process and the activities that designers perform in order to communicate with the design team. In the figure, designers' activities are organized according to the design phases where they mainly occur. Here, the outputs of these activities are also mentioned as 'communication methods'. Figure 1 also summarizes the ideas gathered in this section. Thus, this section will focus on the basic steps of product development processes and later discuss the methods of design communication in relation to the figure.

Product development processes

Designing is considered as a problem-solving activity. According to Roozenburg and Eekels' cycle of design-problem-solving (1995), a solution (i.e., decision about the design) is obtained in five main steps: analysis, synthesis, simulation, evaluation, and decision. This cycle is followed in each and every step of the design processes. That means in each step of the design process, the problems of that stage is analyzed; ideas that form the preliminary solutions are synthesized; suitable solutions are expressed in forms of simulation; the output of the simulation is evaluated; and finally upon the results of the evaluation it is decided on the most optimal solution.

Other methodologies describing the design process (French, 1985; Pahl & Beitz, 1986; VDI 2221, 1987) consist of four main phases (see Figure 1): analysis of the problem, conceptual design, embodiment design, and detailing (see, Cross, 2000 for the detailed comparisons of these design processes):

Analysis of the problem. The nature of the design problem is ill-defined. That is, there is no one definitive formulation of the problem, any problem formulation may embody inconsistencies, formulations of the problem are solution dependent, proposing solutions is a means of understanding the problem, and there is no definitive solution to the problem (Cross, 2000).

Conceptual design. Ideas are generated and assessed in the conceptual design phase. Perhaps, this is the most demanding and mentally exhaustive phase of the design process, because designers consider all aesthetics, ergonomics, emotion, production, and cost related aspects of the product in question and make a synthesis out of them. This is a creative phase in which overall function and important sub-functions of the product are determined. As a result, one or more possible design solutions are generated as concepts. The outcome of this phase has an influential power on the total design of the product. One of the most agreeable concepts is taken further.

Embodiment design. In the next phase (embodiment design), the concept is embodied as the preliminary design of the product and functioning features of the product are defined. This is the phase in which iterations take place in order to refine the design based on technical and economic considerations. The outcome is often drawings of a functioning product on paper and a list of product parts.

Detailing. The definitive design should be detailed in order to safely communicate with the manufacturers and distributors. All technical specifications are done in this phase including materials and parts to be used, dimensions, geometrical shape, etc. The outcome is 'product documents' that encloses technical drawings and instructions for the production assembly, testing, and transport.

These phases described above are not rigidly separated. Overlaps may occur especially between conceptual design and embodiment design phases. Because the goal is not clearly defined in a design problem, the path to solve a design problem—and achieve an optimal solution—happens to be flexible with iterative explorations in each phase of the design process. So, the whole design process can be fed by any alternative solutions at any point in the process—this is especially so if one sees the whole design process as a practical problem solving mechanism.

Design communication

Product design is often conducted by several teams of different disciplines. The design teams simultaneously generate ideas and evaluate them in order to achieve the specified goal. The members of the teams should have a common ground on the design decisions and evaluations. So, the ideas for solutions should be translated equally especially from discipline to discipline. Thus, communication within and between design teams becomes an essential factor in the course of a design process. Well-expressed ideas lead to efficient communication. Efficient communication among the design team speeds up the design process and therefore decreases the cost.

Design communication can be carried out efficiently by the outputs of the *simulation* step in the problem solving cycle. One of the challenges in design communication is to summarize one's ideas and present them to others who might be unfamiliar with the terminology used or the methods applied. To facilitate this, designers perform several activities and use several methods to develop their ideas and communicate them with the design team (see Figure 1). Some of these activities and methods presented in Figure 1 intend to (*i*) visualize the ideas (sketches, drawings), (*ii*) imitate the product-to-be-built (mock-ups, clay modeling, 3D digital models, physical models). (*iii*) present in text the company—or design—values (reports, tables), (*iv*) verbalize the solution-related concerns (discussions, conversations), (*v*) facilitate production (technical drawings), and (*vi*) instruct assembly and use (manuals). These methods of expressions vary in the degree of details depending on the design phase in which they are used (mock-ups might be used in the conceptual design phase, 3D digital models might be used in the embodiment phase, and prototypes might be used in detailing phase).

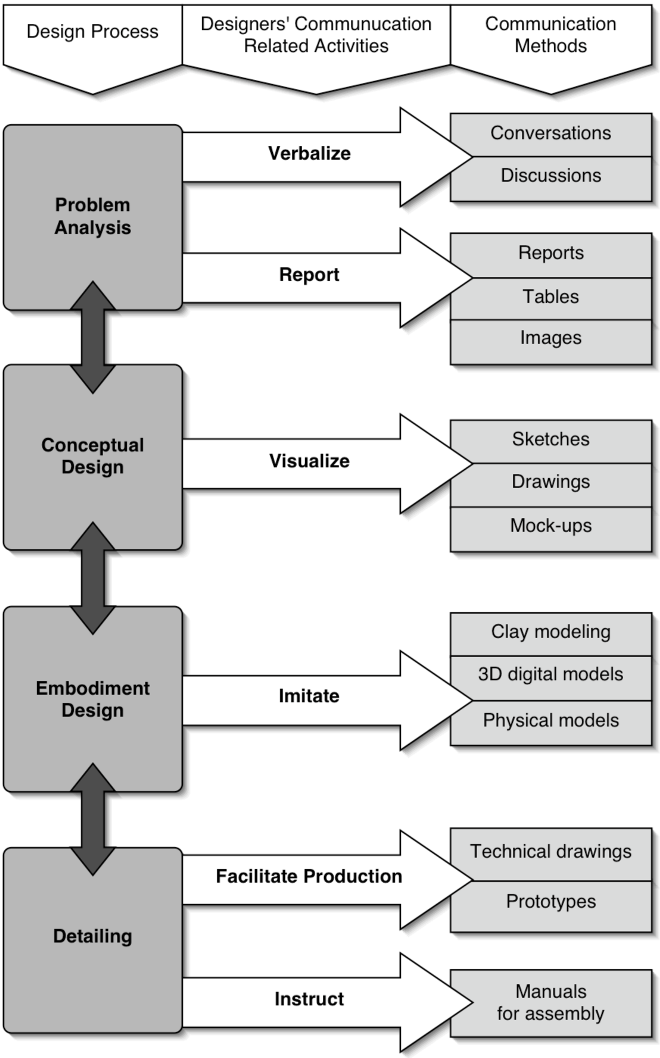


Figure 1. Design process, designers' communication related activities, and communication methods.

Modeling is a very commonly used tool in design communication. Roozenburg & Eekels (1995) discern four main categories of models based on the similarity between the original product and the model: structure, iconic, analogue, and mathematical models. Relevant for sound design are structure and iconic models. *Structure* models can provide a quick first impression of the appearance, functioning, and manufacturing possibilities, and they are often a source of ideas (*p.* 243). The main aim of this type of modeling is the visualization of the qualitative structure of an object. Some examples are sketches, dummies, flow diagrams, etc. *Iconic* models simulate solutions in greater details than structure models. By giving a 3-dimensional overview, they represent the function and the properties (e.g., geometric, thermal, dynamic, etc.) of the original object. Mock-ups, scale models, and prototypes are some of the examples.

Sketching and verbalization are commonly used modes of thinking, reasoning, creating, and discussing. Through *sketching*, ideas are generated and possibilities of product use or functionality are explored. Ferguson (1992) discerns three types of sketching: 'thinking' sketches, 'talking' sketches, and 'prescriptive' sketches. 'Thinking' sketches aid a designer to explore his/her ideas and visualize them. 'Talking' sketches aid a group of designers to explain their ideas and discuss them. 'Prescriptive' sketches aid a designer to communicate to multidisciplinary design team. For (visual) design problems sketching is a powerful tool that aids the visual thinking and expression. By *verbalizing* their ideas, designers convey their concerns and suggestions. In a design team, engineers seem to express themselves well in a verbal conversation, therefore, tend to model their ideas with words (Lloyd et al., 2001); whereas industrial designers often use of graphical representations to express their ideas.

Relying on mere verbal communication may have its own drawbacks. Words may become insufficient to describe, e.g., perceived qualities of the material chosen for a product. The ambiguity in the verbal conversation can be cancelled by, e.g., visualization of an object. Moreover, designing has become an international activity due to cost effectiveness and knowledge share. It is common to see that different—or even the same—design processes are held in different countries. This brings out its own problems because terminologies used in one country might not match in another (Martens & Giragama, 2002) and also because of different cultural backgrounds.

To our knowledge, the above described methods of modeling and communication do not make use of the auditory property of the object-in-design. It is our aim to include 'auditory' models in design communication.

Product sound design

Its short history

For a long time sounds emitted by products have been regarded as noise and therefore as an undesired product feature that should be reduced or eliminated. Noise-control methods have been used to design noise enclosures, isolation systems, and silencers to make the products more acceptable. The main problem with the noise was its loudness. Sound quality control would entail the sound level measurement and comparison of the measurement with the target sound level. For example, when a vacuum cleaner sounded as loud as 78 dB, engineers would design new parts to dampen and isolate the noise which reduced the sound level to, e.g., 70 dB. Because such a method disallows designers to foresee the upcoming problems related to sound, noise-control becomes an independent design process. This results in additional cost due to the extra materials used and man-hours spent. Manufacturers often disfavored noise control because of its costly nature. Noise control methods have not been abandoned—yet there are some application areas for them fitting certain design requirements (Bodden et al., 2002). However, recently design teams started to incorporate sound early in their design decisions rather than solving the problem when it occurs.

In the recent view, product design teams consider sound as one of the inherit features of the sheer product functionality (other product features would be form geometry, material texture, size, weight, etc.). Designers should consider sound as a challenging problem to be solved in the design process and abandon the opinion that sound is as a negative product feature which must be cancelled promptly (Lyon, 2000). Thus, designers should seek ways to explore how to exploit sound to enhance the user experience with products.

Product sound design - now

Sound design is mostly practiced during detailing of the product design process. In the detailing, prototypes are built and trial runs are conducted to simulate the functionality of the original product with the real parts. A functioning prototype naturally creates sound, which is indicative of product's inherit sound. So, this is the phase where the sound quality of the product can be assessed. If the results of the assessment fail to reveal any good correlation to the design requirements of the product, then the sound design needs to be conducted. The parts that fail to produce the desired sound are changed and replaced with another one, and then the sound quality of the product is re-assessed. This iterative process continues until the desired sound is created.

In some cases, concepts and alternatives are created for a sound design. This is done by recording the sound of the prototype and modifying it with the help of a computer. Such a modification is done in a way that it represents the sound of the part that should be replaced with the problematic part. In this stage, suggestions are given with respect to the product parts to be used. After having a few alternatives and the original sound of the prototype, the sound quality assessments can be done on the digital sound files of the suggestions. The modified sound can also be used as a communication tool for the design team to discuss how to proceed further. The most preferred sound is taken further with the requirements for the new prototype to be built.

Product sound quality

The physical character of a sound has psychological correlates (Solomon, 1958; von Bismarck, 1974; Björk, 1985). Similarly, a product sound—depending on its spectral-temporal structure—conveys high-level hedonic attributes rather than evoking only sensory perception/(un)pleasantness. Lageat et al. (2003) links the ‘concrete’ attributes of product sounds (i.e., spectral-temporal properties.) to the product’s hedonic attributes (i.e., pleasant, aggressive, discreet, luxury) suggesting that one can manipulate the ‘concrete’ characteristics of a sound in order to convey hedonic attributes of product through sound. In especially automotive industry, this is a rather exploited area. For example, car manufacturer DaimlerChrysler investigated the degree to which sportiness and sophistication of a car could be represented by the loudness, timbre, and roughness of its engine sound (Letens, 2000). Door sounds of expensive cars are also designed to convey luxury, comfort, and safety.

Similar approaches have been widely studied under the name of product sound quality assessment. Sound quality assessment suggests the adequacy of the sound to the product it belongs to (Blauert & Jekosch, 1997). In other words, the sound should convey the same values as the product. In a framework, Blauert & Jekosch (1997) discuss the process by which users assess product sounds. In this process, the assessment of a sound is done upon auditory perception, and this judgment is continuously fed by cognitive and emotional processes, and by the input from other sensory modalities. This framework also suggests that mere psychophysical measurement of a sound (e.g., sound pressure level or sharpness) does not suffice for determining its psychological effects on users. Other similar approach was posed later by Fog & Pedersen (1999) who point out to the subjectivity of the sound quality measures and unexpectancy of the judgments. In their model user judgments pass through two filters: (a) users’ sensory sensitivity and selectivity towards the product sound, and (b) users’ ‘background, expectations, interest, emotions, and mood’. Lyon’s (2000) approach is similar to the one of Blauert & Jekosch’s (1997). These models intend to map the sound's physical instrumental measure (absolute) to psychoacoustical attributes (objective), and that to user judgments (subjective).

Guski (1997) discusses the methods used to analyze product sound quality and indicates three psychologically relevant aspects: (a) stimulus-response compatibility (reaction time measurements), (b) pleasantness of sounds (questionnaires), and (c) identifiability of sounds (recognition—yes-no—tests and/or verbal descriptions). A questionnaire is the most frequently used method (see, Altinsoy et al., 1998; Lyon, 2000; Lageat et al., 2003) which tests whether the sound conveys the desired attributes of the product. The attributes may represent ergonomics, safety, emotions, hedonics, psychoacoustics, and other attributes depending on product's design. Another method is the analysis of the verbal descriptions. This method is commonly used in recognition test checking the identifiability of product sounds. Interpretations of the verbal descriptions are made to understand the underlying factors of the product sound quality (for a general opinion on the vocabulary listeners use to describe product sounds, refer to Özcan & van Egmond, 2005).

The methods described above can be used either prior to designing sounds to determine the problems with the sound of an existing product or they can be used during prototyping phase of the product to see whether the desired values have been achieved.

Complexity of designing product sounds

Designing sounds for products is a complex process. Below we explain the reasons that contribute to this complexity.

Sound as an indirect result of moving product parts

Because sound is a consequence of moving parts in a product, designing the sound would mean changing the physical properties of the moving parts such as shape, material, and size. When a product is being designed, parts and functionality are determined with respect to the design problem and its requirements for the solutions. It is the interaction of parts and the action involved in the functionality that cause the sound; so, sound design cannot be independent of these aspects. Thus, there should be a good compromise between the design of desired aspects of the product and of the sound.

Consequential outcome of the sound design

Sound design may cause a chain reaction in the design cycle. Hubka and Eder (1988) explain in a framework how design properties are linked to each other and to internal and external properties of the production cycle. Adopting the framework, we can assume that changing a part in the product in order to design the sound may influence the allocated space in the product casing, which would influence the size and weight of the product, and which would affect the packaging design, distribution, and finally the cost of the product.

Physical absence of sound during the design process
In a design process, one cannot talk about the existence of sound during the problem analysis and conceptual design phase. Sound, as a product property, starts to emerge only when the first models of the product is built (embodiment design). Very often only the working prototypes emit sounds that may represent the original sound. Only in the embodiment phase designers consider including sound design in their design conversations. This is not handy as our aim is to include sound design in the conceptual design phase.

Communication about the auditory properties of a product.
It is not clear from the start of the design process what kind of sound stream a product will emit. Designers, therefore, have to rely on their imagination during design meetings to communicate about the psychological effects the sound will have on the users and to predict, in the absence of sound, what needs to be done in terms of sound design. In such cases, the potential product sound needs to be reproduced from memory by retrieving the a priori encoded sounds emitted by similar products or product parts. As the recalled sounds may not clearly represent the potential sound of the product, the judgments would be based on imaginative information that might even be irrelevant to the actual design problem. This may cause misconceptions among the design teams that may lead time and resources loss. However, the early inclusion of the product sound in a design discussion would facilitate the design communication about sounds, and auditory judgments would benefit from it.

Proposed solution

Sound design as a process should run parallel to the main product design process. Design teams should incorporate the sound related problems in their agenda already in the beginning of the product development process and invest effort in it in the subsequent phases. Doing this would prevent later occurring unexpected problems caused by sound. In Figure 2, suggestions for sound design related communication methods during a design process are proposed. This figure works similarly to Figure 1, moreover incorporates designers' sound design related activities and communication methods. Below, these methods are explained in more details following Figure 2.

During the problem analysis, designers should include sound in their *discussions* and *auditorily exemplify* the sound related problem. The examples can be created by recording the sound of the products or by demonstrating the problem with the presence of the working product in question.

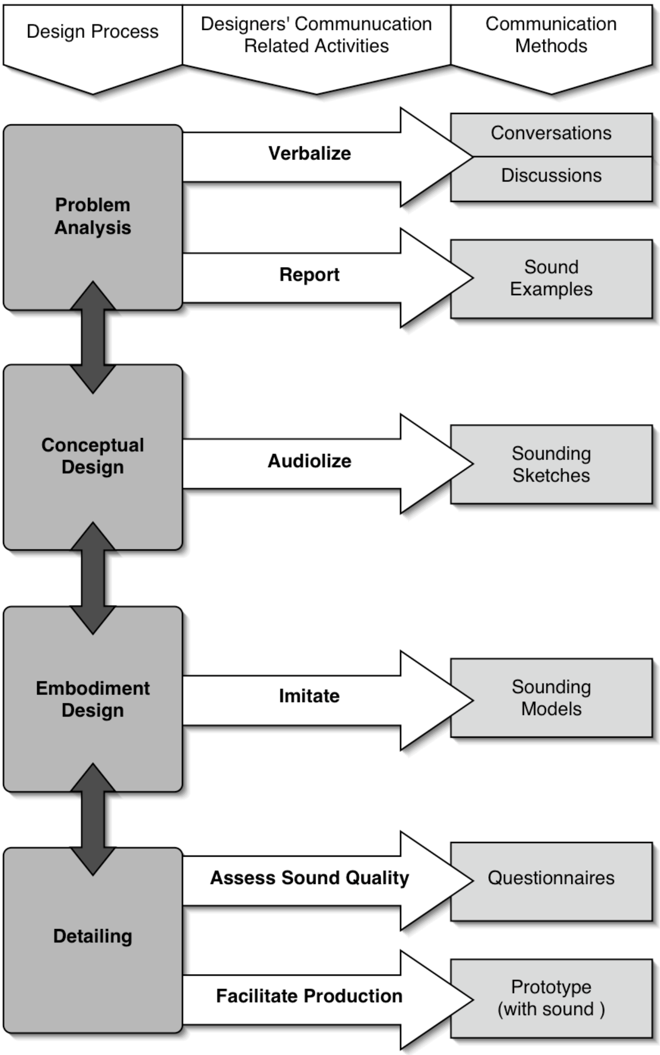


Figure 2. Proposed methods for product sound design related communication.

In conceptual design phase sketching is a common method to explore ideas and visualize them. So, our suggestion is creating *sounding sketches* for sound-related problems in order to 'audiolize' the product sound. Sounding sketches can be recordings of any object that has the potential to represent the sound desired. Or, these sketches can be collections of materials and objects that would exemplify the desired features of the product sound or the product itself. These sound examples might be ambiguous, so, they do not aim to represent the original sound, but they rather represent concepts.

In the embodiment design, the ideas are materialized and parts-to-be-used are determined. Then, models are produced that represent (and imitate) roughly how the product looks or function. So, our second suggestion is creating *sounding models* to simulate the original sound in greater details than sounding sketches. Sounding models can be the composition of sound producing parts. As a communication tool, sounding model summarizes designers' ideas about the proposed sound and makes it easy to discuss the suitability and the feasibility of the proposed solution. However, designers lack a tool to model the sound. Such a tool is under development within our research group and in the next section its functionalities will be briefly explained.

In the detailing phase, a prototype exits to test the functionality of the product. As the sound produced also represents the original sound of the product, sound quality assessments can be done using *questionnaires*. The results of which can be used to determine the final appropriateness of the sound to the product. The *sound of the prototype* can also be used as a reference for the original sound during production.

Tool for creating 'sounding models'

The main problem in simulating sound in the embodiment phase is the lack of sound. If sounds were available, designers would easily be able to create soundscapes (i.e., sound stream of functioning product) and compare the alternatives. Designers may always decompose products to separately record the sound of the parts they need. However, it would be time-consuming and unpractical. Therefore, we are developing a tool that compensates this lack and allows designers to simulate soundscapes.

Sound library

This tool's main feature is the sound library. The library contains previously recorded sounds of product parts. Because the library may not contain all possible parts that exist in the manufacturer or that a designer desires to have, the tool allows designers to manipulate the previously recorded sound. Manipulation is based on changing the physical properties of a sound. By playing with certain parameters, a designer creates the desired sound.

Functionalities of the tool

The tool has three main functionalities that aid a designer to finalize a sound simulation (i.e., sounding model):

Assembly. The tool allows designers to create a functioning product just by using sounds. By positioning the sounds on the timeline according to the order of event occurrence, the simulation is roughly finished.

Design. As described earlier, by manipulating the individual sounds on the specific physical parameters, a designer creates the desired sound.

Evaluation. The tool incorporates algorithms for parameters such as roughness, sharpness, loudness, tonality, and pitch are used to determine the acceptability of the sound on a sensory level.

The tool is very basic and its functionality is based on intuitive actions that allows a designer to explore his/her ideas. So, it can be used by any designer who is novice in designing product sounds. As a first step, we aim design students as prospect users for the tool. We believe that it is important to educate design students as being aware of this rapidly emerging need for product sound design.

Conclusions

Methods aim to help (novice) designers to choose tools suitable for the purpose of their design activity. Methods also guide designers how to use the selected tools, in what condition, and at what stage of the design process. Therefore, we believe that the product sound design related communication methods and the sound modeling tool presented in this paper support this view. The proposed methods enable (sound) designers to systematically tackle the sound design process and to efficiently communicate the sound related design problems/solutions. Thus, the complexity of product sound design will be diminished to a certain extent.

Suggestions for the future

Before designing product sounds, the relation of the sound to the product it belongs to should be examined. Every product exhibits a different character. A proper sound design for one product does not necessarily correspond to a proper sound design of another. It is important to define the problems with product sounds in the context of human behavior, as users have the vote for the acceptability of the product. People exhibit certain behavioral patterns and action tendencies to any object they encounter around them (Plutchik, 1984; Frijda; 1986). Relevant to us is people's reactive behavior to objects. That is, people may *accept*, *reject*, or *ignore* an object depending on the context in which the object is available. When designing sounds, as a first step it is important to determine the psychological effects of the sound on people and their causes. For example, a standard digital alarm clock sound, which is highly present in an environment, can be found unpleasant. People may *reject* this sound due to the sound quality—high-pitched, sharp, loud, etc.—however, considering the function of the product they may *accept* it. Or in another example, a loud sound may be *rejected* for a city car because it causes noise-pollution, *accepted* for a sports car because it indicates power, or *ignored* for an old classic car because it only signifies old-fashioned technology. These behavioral concerns could form the guidelines for the type of sound to be designed.

Acknowledgments

The tool is being developed by Erik Huijs as a part of his graduation assignment for Faculty of Industrial Design of Delft University of Technology under the supervision of the authors, Piet Westendorp, and Paul Mijksenaar. Therefore, we would like to thank them for their contribution for building up the concept of the tool. We are also grateful for Norbert Roozenburg and Petra Badke-Schaub for their valuable comments on the product sound design related methodological issues. Finally, we would like to acknowledge the useful comments of an anonymous reviewer.

References

Alessi, A. (2006). in *Le Mag*, Euronews. 21 April 2006.

Altinsoy, E., Kanca, G., & Belek, H. T. (1998). Ürün gelistirmede ses kalitesi yöntemi ve elektrikli süpürgeler üzerinde bir uygulama. Proceedings of 4. Ulusal Akustik Kongresi, Antalya, Türkiye, September (Turkish).

Björk, E. A. (1985). The perceived quality of natural sounds. *Acustica, 57*, 185-188.

Blauert, J., & Jekosch, U. (1997). Sound quality evaluation: A multi layered problem. *Acustica, 83*, 747-753.

Bodden, M. (1997). Instrumentation for sound quality evaluation. *Acustica, 83,* 775-783.

Bodden, M., & Iglseder, H. (2002). Active sound design: Vacuum cleaner. Forum Acousticum, Seville, Spain.

Cross, N. (2000). *Engineering design methods: Strategies for product design* (third edition), Chichester: John Wiley and Sons.

Desmet, P. M. A. (2002). *Designing emotions*. Doctoral Dissertation, Delft University of Technology, Delft, The Netherlands.

Ferguson, E. S. (1992). *Engineer and the mind's eye: Sketching and visual cognition*. Cambridge: MIT Press.

Fog, L., & Pedersen, T. (1999). Tools for product optimization. Human Centered Processes 10th Mini EURO Conference. Brest, France.

French, M. J. (1985). *Conceptual design for engineers*. London: The Design Council.

Frijda, N. H. (1986). *The emotions*. Cambridge: Cambridge University Press.

Guski, R. (1997). Psychological methods for evaluating sound quality and assessing acoustic Information. *Acustica*, *83*, 765 -774.

Hubka, V., & Eder , W. E. (1988). *Theory of technical systems: A total concept theory for engineering design*. Berlin: Springer.

Lageat, T., Czellar, S., & Laurent, G (2003). Engineering hedonic attributes to generate perceptions of luxury: Consumer perception of an everyday sound. *Marketing Letters, 14*(2), 97-109.

Letens, U. (2000). Exemplarische darstellung einer subjektiven beurteilung von PKW-fahrgeräuschen. Paper presentes at the 26. Deutsche Jahrestagung für Akustik, DAGA, Deutschland.

Lloyd, P. A., & Busby, J. S. (2001). Softening up the facts: Engineers in design meetings. *Design Issues, 17*, 67-82.

Lyon, R. H. (2000). Designing for product sound quality. New York: Marcel Dekker.

Martens, W. L., & Giragama, C. N. W. (2002). Relating multilingual semantic scales to a common timbre space. Paper presented at the 113th Convention of the Audio Engineering Society, USA.

Özcan, E., & van Egmond, R. (2005). Characterizing descriptions of product sounds*. Proceedings of the International Conference on Auditory Display, Ireland, 11,* 55-60.

Özcan, E., van Egmond, R., & Jacobs, J. (submitted). Bases for categorization and identification of product sounds.

Pahl, G., & Beitz, W. (1986). *Engineering design: A systematic approach* (second edition) London: Springer.

Plutchik, R. (1984). Emotions: A general psychoevolutionary theory. In K. R. Scherer & P. Ekman (Eds), *Approaches to Emotion* (pp. 197-219). Hillsdale: Erlbaum.

Roozenburg, N.F.M., & Eekels, J. (1995). *Product design: Fundamentals and methods.* Chichester: Wiley.

Solomon, L. N. (1958). Semantic approach to the perception of complex sounds. *The Journal of the Psychoacoustical Society of America, 30*(5), 421-425.

Västfjäll, D. (2002). Emotion in product sound design. *Proceedings of Design Sonore*. Paris, France.

Västfjäll, D., Gulbol, M. A., Kleiner, M., & Gärling, T. (2002). Affective evaluations of and reactions to exterior and interior vehicle auditory quality. *Journal of Sound and Vibration, 255*(3), 501-518.

Västfjäll, D., Kleiner, M & Gärling, T. (2003). Affective reactions to interior aircraft sound quality. *Acustica, 89*, 693-701.

VDI Guideline 2221 (1987*). Systematic approach to the design of technical systems and products*. Düsseldorf: VDI (edited by K.M. Wallace).

Von Bismarck, G. (1974). Timbre of steady sounds: A factorial investigation of its verbal attributes*. Acustica, 30*, 146-159.

This chapter is based on the paper:

Özcan, E., & van Egmond, R. (2008). ***Product Sound Design: An Inter-Disciplinary Approach?*** **Paper to be presented at the 4th International Conference on Design Research, Sheffield, UK.**

Abstract

The practice of product sound design is relatively new within the field of product development. Consequently, the responsibilities and the role of a (sound) designer are not very clear. However, practice shows that various disciplines such as design engineering, acoustics, psychoacoustics, psychology, and musicology contribute to the improvement of product sounds. We propose that sound design should be conducted by experts who have knowledge in the afore-mentioned fields. In other words, we suggest that product sound design should be an independent field that encompasses an inter-disciplinary approach.

CHAPTER 8

⁂

PRODUCT SOUND DESIGN: AN INTER-DISCIPLINARY APPROACH?

Our daily interaction and experience with the sounds that products emit are various. One could have a desire for a car because of its sophisticated door and engine sound, or one may despise an alarm clock sound because it is too loud and too sharp. Using a vacuum cleaner may be too uncomfortable to one's ears, however the happy bell of a microwave oven may be the most expected sign for a late dinner. These examples illustrate the influence of product sounds on our reasoning, on our emotional state, on our purchasing decisions, and on our expectations regarding the product and its functionality.

Studies regarding product sound design and perception have also confirmed the complimentary role of auditory experience on how people perceive and respond to products (Lageat, Czellar, & Laurent, 2003; Vastfjall, Kleiner, & Garling, 2003; van Egmond, 2008). That is, a well-designed sound enhances the product experience on ergonomic and hedonic levels. Conversely, unsatisfactory auditory experience will negatively influence one's emotional responses to and conscious judgments on a product. Therefore, in the last decade, more attention has been dedicated to improve the quality of product sounds and consequently the product experience (Lyon, 2000; Özcan & van Egmond, 2006; van Egmond, 2008).

Although designing sounds for products have become a rather acknowledged practice within the field of product development, the task of a designer with respect to sound design is not very clear. In an average sound design task, it is expected that the sound of a product is adequate to the product it belongs to (Blauert & Jekosch, 1997). For example, a kitchen extractor fan should sound 'powerful, yet inconspicuous'. However, for designers, achieving such a goal is not very straightforward. Designing product sounds entails an iterative exchange of expertise from various disciplines that are functionally different. In principle, designing sounds for products requires manipulation of the structural and material configuration of products—because a product sound is a consequence of moving product parts. Primarily, an acoustical analysis is required to determine the physical character of the sound (i.e., spectral-temporal structure), which can then also be used for sound simulations (Lyon, 2001; Susini, McAdams, Winsberg, Perry, Viellard, & Rodet). A psycho-acoustical analysis reveals people's sensorial reactions to a sound in terms of pleasantness or comfort (Zwicker & Fastl, 1990). Furthermore, semantic associations of the created sounds need to be tested for the adequacy of the sound to the product (Blauert & Jekosch, 1997; Guski, 1997,). In some cases of sound design, musical knowledge is required to compose somewhat musical sounds (e.g., mobile phone ring tones, alarm clocks) (Schimmel, 2001). Thus, the fields of acoustics, psycho-acoustics, engineering, psychology, and musicology contribute to the improvement of the sound at different stages of a sound design process. The multi-disciplinary nature of product sound design makes the design practice too complicated for an average designer / design engineer. Therefore, the tasks regarding the sound design should be separated from the tasks of design engineers.

We propose that sound design, instead of being a multi-disciplinary practice that requires the simultaneous involvement of various experts, should be considered as an inter-disciplinary practice that is conducted by experts who have knowledge in the afore-mentioned fields. Thus, in this paper, we will focus on the contribution of various disciplines to product sound design. Furthermore, the responsibilities of a *sound* designer will be discussed and the plausibility of product sound design as an independent field will be argued.

Product sounds

Two types of product sounds exist: *consequential* sounds and *intentional* sounds. *Consequential* sounds are emitted by products as a result of their functioning. For example, the sound of a hairdryer, vacuum cleaner, washing machines, etc. are considered to be consequential sounds. Such products contain multiple sound producing parts such as running engines, rotating gears or fans, bouncing springs, pumping water, blowing air. The formation of the product sound is dependent on the type of action and the type of source in action. For example, if the product is electrically operated, it probably contains an engine and a gearbox. Attached to them may a fan that has to rotate or blades that have to move and cut. A rotating fan may be used to blow or suck air. Moreover, the material, size, and the geometry of the product part also contribute to how the sound is formed. Consequential sounds are often informative about the product functioning cycle and listeners cannot intervene their occurrence. *Intentional* sounds are designed, implemented, and put by a sound engineer. Microwave oven finish bells, alarm clocks, oven setting feedback sounds are some of the examples. They are mostly digital and somewhat musical sounds often used in user interfaces. Such sounds are abstract by nature; however, listeners learn to attribute meaning to them as they are mostly designed to convey certain messages. Listeners also feel obligated to attend to intentional sounds due to their communicative nature.

Furthermore, product sounds can be discerned into six perceptually distinguishable sound categories (see Chapter 1). These categories are air, alarm, cyclic, impact, liquid, and mechanical sounds. Sounds in these categories vary in their spectral-temporal composition, material interactions that cause sound, and conceptual associations. In addition, the perceived character of a sound can be dependent both on perceptual and cognitive factors (Özcan & van Egmond, 2007; see also Chapter1).

Defining the field of product sound design

Why design product sounds?

Design problems concerning product sounds are situation based. Although silence is preferred for some products (e.g., computer fans, dishwashers), the presence of a sound is almost compulsive when it comes to cars, espresso machines, or alarm clocks. For example, a computer is expected to be silent because it is a heavy-use domestic appliance which should function inconspicuously. However, the experience of a car ride may be complete with the proper auditory feedback that is responsive to certain user actions (e.g., acceleration or breaking) or that is suitable to the character of the car (e.g., sports car). Products such as alarm clocks exist merely because of their auditory function. Furthermore, because sound is a consequence of a functioning product, its presence can be complementary to user expectations regarding the product. For example, it may be the sound of an espresso machine that prepares a person to a tasteful Italian coffee. In summary, comfort, ergonomic use, functionality, or hedonic values may constitute the main reasons to design the sound of a product. Nevertheless, whatever the reason is, the main concern regarding product sound design is the suitability of the sound to the concept of the product (Blauert & Jekosch, 1997; Özcan & van Egmond, 2006).

Sound design within industry

Designed sound often indicates sophistication in the engineering of the product, thus increases the perceived value of the product. Especially automotive industry has dealt with the improvement of the sound of their products. To our knowledge, they have specifically designed the sound of the door-closing (Kuwano, Fastl, Namba, Nakamura, & Uchida, 2006), engine (Letens, 2002), gearbox (Bodden & Heinrichs, 1999) and tested the user responses to the changes in the sound quality (Blauert & Jekosch, 1993; Bodden, 1993; Bisping, 1997). Sound design can also be found in other product domains such as crunchiness of a crisp or the softness of the plastic bottle of a fabric softener are all designed to complement the product experience. Although there is an increasing interest in the sound design of domestic appliances, the sound design of the domestic appliances has been mostly restricted to noise closures and diminishing the loudness of domestic appliances (Lyon, 2000). In domestic appliances, added sounds are often used to communicate abstract meanings or provide feedbacks. The keystroke tones in mobile phones, the bell of the microwave oven, and the click of the mouse are some examples.

Available tools and methods

Both the industry and the academia are interested to develop tools and methods for the design of product sounds. Industry reveals only little information regarding the tools and methods used for the sound design practice. However, a well-known method to judge the suitability of the sound to the product is the sound quality assessment (Blauert & Jekosch, 1997). For that, a questionnaire is used that contains a list of adjectives that have potential to describe the sound in development. As a result, product developers are able to test upfront psychological effects of the designed sounds (see e.g., Kuwano et al. 2006, Letens, 2000).

Other methods have been developed to predict the perceptual space for the sound in development. For example, listeners' preference for noisy appliances could be predicted using psycho-acoustical data such as loudness, harmonicity, and noisiness (Susini et al, 2004). When diagnosing fault in product parts, acoustical measurements can be helpful (Benko et al.). Bodden (1997) suggests that such predictions and the auditory analysis of the product sound should be done considering the users and the context of use.

The application of product sound design

Sound is an integral property of the product. Any changes on sound require changes in the product. Thus, the application of product sound design is a part of the main product development process and should run in parallel to it. An iterative problem analysis and solution is conducted regarding the source of the sound (i.e., product and its parts). Özcan and van Egmond (2006) have suggested that the process of sound design is very similar to those processes of product development proposed by Roozenburg and Eekels (1995).

Similarly, the process of product sound design consists of four main phases: problem analysis, conceptual design, embodiment design, and detailing (see Figure 1). In *problem analysis phase*, designers verbally discuss and auditorily exemplify the sound related problem. The examples can be created by recording the sound of the products or by demonstrating the problem with the presence of the working product in question. In *conceptual design phase*, designers auditorily sketch their conceptual ideas. Sounding sketches can be recordings of any object that has the potential to represent the sound desired. These sound examples may be ambiguous, and do not aim to represent the original sound. In *the embodiment design*, the ideas are materialized and parts-to-be-used are determined. Then, sounding models are produced that represent (and imitate) roughly how the product functions and will sound accordingly. As a communication tool, sounding model summarizes designers' ideas about the proposed sound and makes it easy to discuss the suitability and the feasibility of the proposed solution. In *detailing phase*, a prototype exits to test the functionality of the product. As the sound produced also represents the original sound of the product, sound quality assessments can be done using questionnaires. The results of which can be used to determine the final appropriateness of the sound to the product.

Bodden (1997) has suggested that for good auditory analysis, equipment specific to product sound analysis is required. Signal acquisition should be done carefully by using multi-channel recording methods to capture more auditory information. Later, basic signal analysis methods (e.g., adopted from Zwicker & Fastl, 1993) are applied to understand the acoustic nature of the sound (i.e. spectral and temporal composition of the sound). Relevant modeling and editing techniques are used to simulate the desired sound. However, results work the best when sound and source are coupled for the sound quality evaluation.

Analyzing the acoustic property of the sound and determining the problem is the first step. Sound simulations already suggest the desired output of the design process. However, the next critical step is the materialization of the ideas. That is, the design team needs to formulate what product part needs to be changed or replaced, what product part actions need to be calibrated, and how the order of events should occur in order to offer the desired output. This may be an iterative process which requires high technical skills on components, structures, and assembly for the well-tuning of the sound (Lyon, 2000).

Major studies in the field of product sound design all agree on the psychological effect of sound on users (Blauert & Jekosch, 1997; Bodden, 1997; Lyon, 2000; Lyon 2003 Özcan & van Egmond, 2006; van Egmond, 2008). It is the user that determines the adequacy of the sound to the product. Therefore, especially in the last phase, but preferably throughout the whole design process, user input needs to be considered.

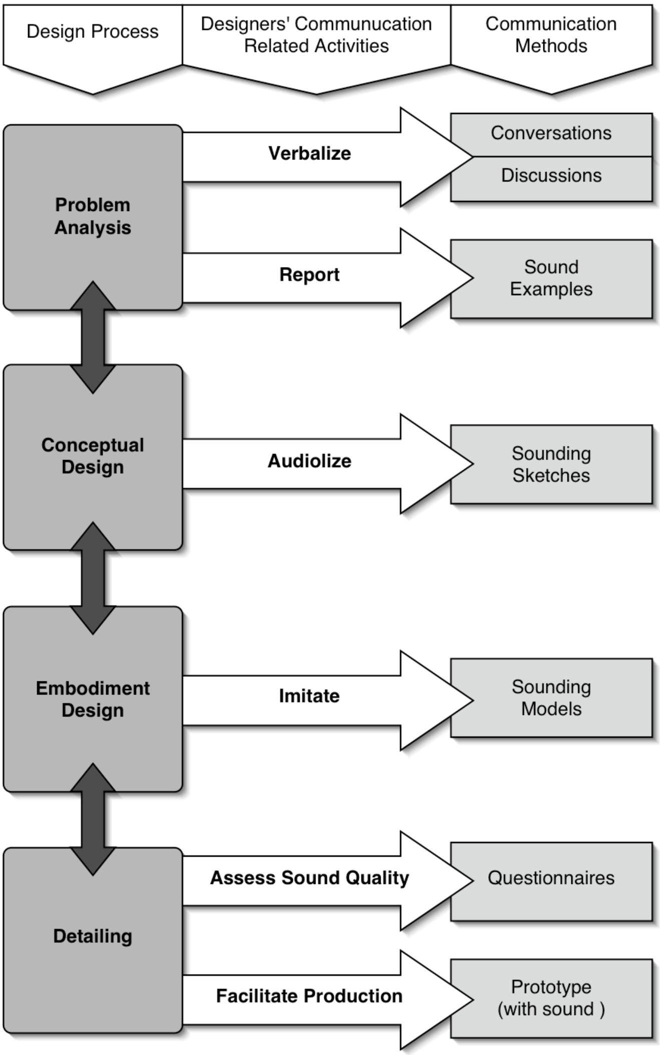


Figure 1. Proposed methods for product sound design related communication.

The use of questionnaires is one way of verifying the semantic and conceptual relation between the sound and the product. However, theoretical studies provide insight into conceptual network regarding product sounds and cognitive processes that underlie such network (Özcan & van Egmond, 2007; see also Chapter 1). This means that design teams could incorporate such knowledge into auditory sketching and conceptual design of the sound.

Disciplines contributing to product sound design

Any design process has the potential be multi-disciplinary. Experts from different fields may contribute to a design activity depending on the task and requirements. For sound design, three indispensable disciples provide knowledge: *acoustics*, *engineering*, and *psychology*. A sound design task cannot be completed in the absence of one of these disciplines. Figure 2 demonstrates how knowledge from these disciplines feed the sound design process and results in the main solution provided for the sound problem of the product. In the following paragraphs we will explain the individual contribution of these different fields of expertise.

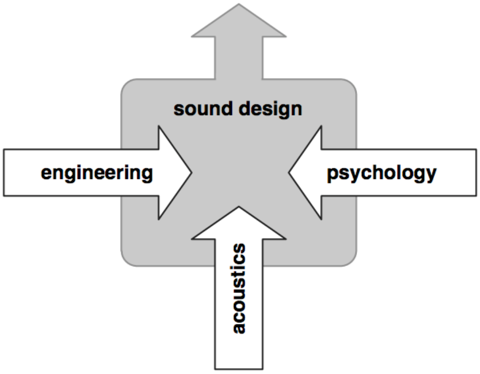


Figure 2. Main disciplines contributing to product sound design activity.

Acoustics

Acoustics is the science that focuses on the sound phenomenon. It covers basic physical principles related to sound propagation and mathematical and physical models of sound measurement. Therefore, the medium in and through which sound travels, reflecting and vibrating surfaces, speed of sound, and other physical characteristics of sound such as sound pressure, wavelength and frequency are the topics of interest for the field of acoustics.

Sound occurs a consequence of the energy release caused by objects in action. Although, the sound source and action determine the physical quality of the sound, *acoustics* does not investigate the source as a whole but the physical properties of the source such as the interacting materials, weight, size, geometry of the objects Furthermore, sound propagates over time because it is the result of time-dependent dynamic events. That is, the physical character (i.e., spectral-temporal composition) of a sound changes over time depending on the type of actions and sound sources. For example, a musical instrument produces a structured sound (due to the harmonic partials and temporal pattern). A shaver produces a noisy sound because it contains multiple sound producing events each creating different harmonic partials and occurring at different time frames causing temporal irregularity.

The field of acoustics provides techniques to analyze and simulate sound. First, basic acoustic terminology consists of *frequency* (variation rate in the air pressure), *amplitude* (magnitude of sound travel) and *intensity* (loudness). Frequency content of a sound and the intensity variations in time are visualized by a spectrogram. Furthermore, a sound wave represents the temporal tendency of sound propagation and the sound pressure over time. Thus, the spectral-temporal composition of a sound event can be visually analyzed and consequences of certain events can be precisely detected. Moreover, various sound modeling techniques have been developed in the field of acoustics. With the available computer technology, it has been possible to simulate sounding objects that are perceptually convincing (Cook, 2002; Pedersini, Sarti, & Tubara, 2000; Petrausch, Escolano, & Rabenstein, 2005; Rocchesso, Bresin, & Fernstrom, 2003).

When designing product sounds, understanding the acoustic nature of the sound event is compulsory. Acoustic analysis of the sound can be first done during problem analysis phase and can recursively occur until the problem has been defined. Furthermore, sound simulation can also be necessary to test upfront the perceptual effects of the desired sound.

Engineering

Engineering is the discipline through which abstract scientific knowledge takes on an applied nature. Regarding product sound design, especially mechanical engineering, electric-electronics engineering, and material sciences provide knowledge. Because sound is a consequence of interacting materials, relevant engineering disciplines deal with sound indirectly and rather focus on manipulative aspects of products. Therefore, various product parts, mechanisms, assembly structure, material interactions, the order of events occurring can all be engineered depending on the design requirements of the product and its sound.

The main focus in product engineering is on the functionality of the product. Thus, suggested alterations that are necessary to improve the product sound can only be done if it does not compromise the main functionality of the product or product parts. Engineers should have satisfactory knowledge on physics and mathematics, therefore are able to calculate the energy release as sound or as vibration. As a result, they can provide solutions in the form of noise closures or sound dampening techniques.

Furthermore, the discipline of engineering provides various tools and methods to embody conceptual ideas and solutions to problems. Engineers and designers are well-supported on modeling, testing, and prototyping (Cross, 2000; Hubka & Eder, 1988; Roozenburg & Eekels, 1995).

Psychology

So far, the contributing disciplines have dealt with the physical aspect of sound and the object causing the sound (i.e., product). However, any sound has psychological correlates which may be on a semantic level or an emotional level (von Bismarck, 1974; Kendall & Carterette, 1995; van Egmond, 2004). Upon hearing listeners' main reaction to a sound is to interpret it. Such interpretations may sometimes be abstract, but they often refer to the source of the sound and the action, such as, crashing car or car passing by (Fabiani, Kazmerski, Cycowicz, & Friedman, 1996; Marcell, Borella, Greene, Kerr, & Rogers, 2000). Many experimental studies have also indicated that just by hearing listeners can describe the material, size, and shape of the sound (Hermes, 1998; Lakatos, McAdams, & Causse, 1997) Listeners are able to follow the changes in the spectral-temporal structure of the sound and perceive it as auditory events or sometimes as auditory objects (Kubovy & van Valkenburg, 2004; Yost, 1990).

The conceptual network for product sounds consist of associations on different levels (see Chapter 1). Source and action descriptions occur the most and followed by locations in which products are used the most (e.g., bathroom, kitchen), basic emotions (e.g., pleasant-unpleasant), psychoacoustical judgments (e.g., sharp, loud, rough). In addition, source properties can also be identified (e.g., interacting materials or sizes of the products). Furthermore, listeners can associate the product sounds to more abstract concepts such as danger. We have also shown that semantic or emotional judgments are sound type dependent (see Chapter 1). For example, alarm sounds are described mostly by abstract meanings such as 'wake up call'; however, impact sounds are described mostly by action and interacting materials.

These conceptual associations of sound indicate that a fittingness of the sound to the product or to the environment in which the sound occurs is judged. Therefore, a design team cannot overlook the cognitive and emotional consequences of the sound. In various stages of design, user input needs to be considered.

Hybrid disciplines: Psycho-acoustics and musicology

Above we discussed the major disciplines contributing to sound design. However, some hybrid disciplines also contribute such as psycho-acoustics and musicology.

Psychoaoustics deals with the basic psychological reactions to the acoustic event. Often the following parameters are used to observe listeners: sharpness (high frequency content), roughness (fluctuation speed of the frequency and amplitude modulation), loudness (sound intensity), and tonalness (amount of noise in a sound). Although these parameters are supposed to be subjective, still a general conclusion has been made in the past regarding the threshold and limits of human sensation to sounds. Therefore, psychoacoustical algorithms have been presented to measure the above-mentioned perceived characters of sound (Zwicker & Fastl, 1990). These algorithms are used to measure the sound's perceptual quality and predict listeners' tolerance to sounds. Thus, they are predictive of sensory (un)pleasantness.

The contribution of **musicology** to product sound design comes when alarm-like synthesized sounds need to be designed. Composing music requires knowledge on theories about musical structures and compositions, tools to create harmonic and rhythmic sounds.

Responsibilities of a sound designer

To sum up, a sound designer needs to have knowledge and skills on three major disciplines (engineering, acoustics, and psychoacoustics) and also on hybrid disciplines such as musicology and psychoacoustics (see Figure 3). A sound designer is primarily an engineer who is able to manipulate the construction of a product and is skillful in applying physical and mathematical knowledge in order to analyze and model product structure while considering the consequences in terms of sound.

However, such an engineer should be able to interpret the physics of sound per se. Skills on acoustic analyses and ability to simulate sound are necessary. Furthermore, a sound designer should be able to link the structural properties of a sound to its acoustical composition. In addition, musical knowledge on how to compose synthesized sounds may be required.

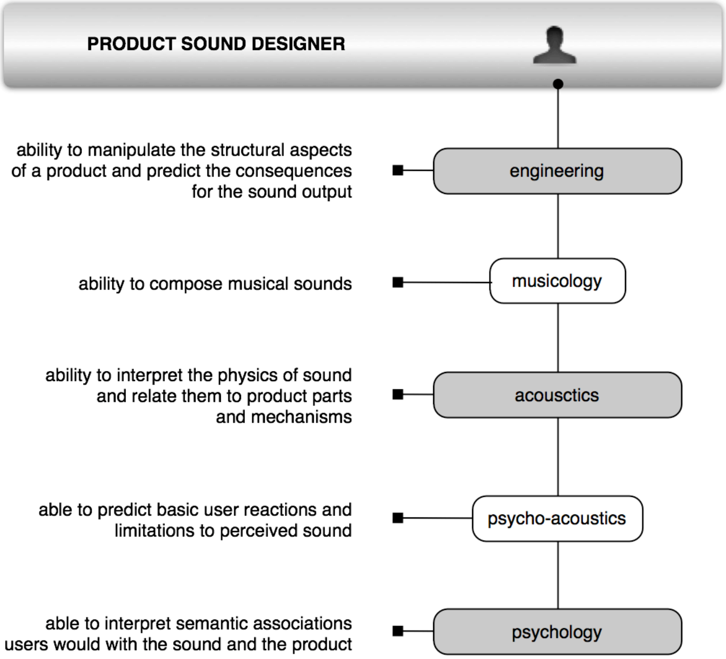


Figure 3. Professional domain of a sound designer.

Furthermore, an engineer solving a product sound problem not only considers the physical aspects of sound and the sound source but also its psychological correlates. It is ultimately the user's vote that counts when judging whether the sound fits the product, its functionality and the context of use. Knowledge on psycho-acoustical analyses is required to predict the first user reactions only to sound. Later, semantic analyses need to be conducted with potential users to make sure the sound design is complete and appropriate to the product.

Conclusions

Is product sound design an emerging discipline?

Sound design practice has long been conducted by a team of designers and engineers who are individually experts in acoustics, engineering, and psychology. If at all a sound designer existed in a design team, this person was more a mediator who made sure that the team members communicated well with regard to the product and its sound and the project was well completed with the contribution from the above-mentioned disciplines. The contributions of the experts from different disciplines made the sound design task a multi-disciplinary task. However, product sound design consists of various recursive tasks. Thus, the sound design process often suffers from communication related problems and recursiveness of such a multi-disciplinary task may hinder the speed and proper application of the solutions. Therefore, instead of having experts from different fields designing the sound of a product, we suggest that a sound designer who has knowledge mainly in engineering and other supporting fields (acoustics and psychology) should take over the sound design task. Embedding the knowledge from different disciplines in one would make the sound design process an inter-disciplinary process rather than multi-disciplinary.

Considering the interest from both the industry and the academia, the tools and methods design specially for sound analysis and design, the body of knowledge that is required to conduct a simple sound design task, we can conclude that product sound design is definitely an emerging discipline. However, yet much needs to be done in order to for this newborn discipline to settle. One main suggestion would be to educate design students on this topic. Schools of industrial design and design engineering should start to include sound design in their curriculum. Furthermore, companies that manufacture products and product ideas could pay more attention the sound design task, consider it as part of the main design problem, and recruit experts—that is sound designers—who are knowledgeable in this field.

References

Bisping, R. (1997). Car interior sound quality: Experimental analysis by synthesis. *Acustica, 83*(5), 813-818.

Bodden M. (1997) "Instrumentation for sound quality evaluation" Acta Acustica united with Acustica 83 775 – 783

Bodden, M. & Heinrichs, R. (1999) Analysis of the time structure of gear rattle. Proceedings of the. Internoise 99, Fort Lauderdale, USA.

Blauert, J., & Jekosch, U. (1997). Sound quality evaluation - A multi-layered problem. *Acustica, 83*, 747 - 753.

Cook, P. R. (2002). *Real Sound Synthesis for Interactive Applications*. Natick: Peters.

Cross, N. (2000) *Engineering Design Methods: Strategies for Product Design* (third edition), Chichester: John Wiley and Sons Ltd.

Fabiani, M., Kazmerski, V. A., Cycowicz, Y. M., & Friedman, D. (1996). Naming norms for brief environmental sounds: Effects of age and dementia. *Psychophysiology, 33*(4), 462-475.

Guski, R. (1997). Psychological methods for evaluating sound quality and assessing acoustic information. *Acustica, 83*(5), 765 - 774.

Hermes, D. J. (1998). *Auditory Material Perception* (Annual Progress Report No. 33): IPO.

Hubka, V. and Eder, W. E. (1988) *Theory of Technical Systems: A Total Concept Theory for Engineering*

Design. Berlin: Springer.

Kendall, R. A., & Carterette, E. C. (1993). Verbal attributes of simultaneous wind instrument timbres: von Bismarck's adjectives. *Music Perception, 10*(4), 445-468.

Kubovy, M. & Van Valkenburg, D. (2001). Auditory and visual objects. *Cognition*. *80*(1–2), 97-126.

Kuwano, S., Fastl, H., Namba, S., Nakamura, S., & Uchida, H. (2006) Quality of door sounds of passenger cars. Acoustical Science and Technology, 27 (5).

Lakatos, S., McAdams, S., & Causse, R. (1997). The representation of auditory source characteristics: Simple geometric form. *Perception & Psychophysics, 59*(8), 1180-1190.

Lageat, T., Czellar, S., & Laurent, G. (2003). Engineering hedonic attributes to generate perceptions of luxury: Consumer perception of an everyday sound. *Marketing Letters, 14*(2), 97-109.

Lyon, R. H. (2000). *Designing for Product Sound Quality*. New York: Marcel Dekker, Inc.

Marcell, M. E., Borella, D., Greene, M., Kerr, E., & Rogers, S. (2000). Confrontation naming of environmental sounds. *Journal of Clinical and Experimental Neuropsychology, 22*(6), 830-864.

Özcan, E., & van Egmond, R. (2007). Memory for product sounds: The effect of sound and label type. *Acta Psychologica, 126*(3), 196-215.

Özcan, E., & van Egmond, R. (2006). *Product sound design and application: An overview. Proceedings of the International Conference on Design and Emotion, Sweden*, *5,* 1-19.

Özcan, E., van Egmond, R., & Jacobs, J. (submitted). Bases for categorization and identification of product sounds.

Pedersini, F., Sarti, A., & Tubara, S. (2000). Object-based sound synthesis for virtual environments - Using musical acoustics. *Ieee Signal Processing Magazine, 17*(6), 37-51.

Petrausch, S., Escolano, J., & Rabenstein, R. (2005). *A General approach to block-based physical modeling with mixed modeling strategies for digital sound synthesis.* Paper presented at the IEEE International Conference on Acoustics, Speech, and Signal Processing, Pennsylvania, USA.

Roozenburg, N.F.M. And Eekels, J. (1995) *Product design: Fundamentals and methods.* Chichester: Wiley.

Rocchesso, D., Bresin, R., & Fernstrom, M. (2003). Sounding Objects. *IEEE Multimedia, 10*(2), 42-52.

Schimmel, O. (2001). Auditory displays. *In* K. Baumann, Thomas, B (Ed.), *User Interface Design for Electronic Appliances* (pp. 253 – 267). Bristol PA, USA: Taylor & Francis.

Susini, P., McAdams, S., Winsberg, S., Perry, I., Viellard, S., & Rodet, X. (2004). Characterizing the sound quality of air-conditioning noise. *Applied Acoustics, 65*(8), 763-790.

Van Egmond, R. (2008). The experience of product sounds. In H. N. J. Schifferstein & P. Hekkert (Eds.), *Product experience*. Amsterdam: Elsevier.

Van Egmond, R. (2004). Emotional experience of frequency modulated sounds: implications for the design of alarm sounds. In D. de Waard, K. A. Brookhuis & C. M. Weikert (Eds.), *Human factors in design* (pp. 345-356). Maastricht: Shaker Publishing.

Vastfjall, D., Kleiner, M., & Garling, T. (2003). Affective reactions to and preference for combinations of interior aircraft sound and vibration. *International Journal of Aviation Psychology, 13*(1), 33-47.

Von Bismarck, G. (1974). Timbre of steady sounds: A factorial investigation of its verbal attributes. *Acustica, 30*, 146 - 159.

Yost, W. A. (1991). Auditory image perception and analysis: The basis for hearing. *Hearing Research, 56*(1-2), 8-18.

Zwicker, E., & Fastl, H. (1990). *Psychoacoustics: Facts and models*. Berlin: Springer.

IMPLICATIONS

⁂

The findings of this thesis demonstrate that the psychological effects of product sounds are undeniable. These effects should be considered at all times for the design, marketing, sales, and the use of the product. What do all these findings mean when it comes to designing product sounds? How can designers use this new body of knowledge? In this section of the thesis, I will relate the empirical findings to the application of product sound design and present some guidelines for (sound) designers. Future suggestions and limitations of the present knowledge are included in this section.

Defining product sound design

With the new knowledge provided by this thesis, product sound design deserves a new definition. In this definition, first, psychology needs to be included in the list of disciplines contributing to the knowledge concerning product sounds. Understanding the psychological effects of product sounds will provide human-centered engineering solutions to product sound design. Secondly, sound should be considered as an integral property of the product. Conceptual associations elicited by the product eventually reflect on the meaning derived from or attached to product sounds. Consequently, a conceptual congruency will often be required between a product and its sound.

There are many motivations to design product sounds. Imagine the sound of a scooter. What everybody absolutely hears is the high-pitched, rather rough, and irregular sound of a 2-stroke engine. Surprisingly, not everybody reacts to this sound similarly. A rider on a scooter will primarily use the sound as feedback to see whether the engine is responding well to acceleration or deceleration. However, he can also produce a wilder sound to draw attention of a nice girl waiting to cross the street. After all he thinks the wild sound will go well with the flame patterns on his scooter. While this girl may not appreciate this wild sound, she will definitely respond to the warning quality of it. She will try to determine whether this motor vehicle is approaching too fast to the zebra crossing. A retired man living on this street may be too weary of these frequently occurring scooter sounds; whereas his hungry next-door neighbour awaits a similar sound because he has just ordered pizza! This example, demonstrates that a sound's function in a product varies and meanings associated to the sound depend on the person.

In the following paragraphs, I will elaborate more on these motivations and give examples using other product sounds.

Auditory ergonomics. Sounds can be informative about a product's function or working cycle. The warning sound of a car seat belt should inform people correctly about its purpose ('danger if not fastened'). A washing machine's wash-cycle sound should be recognizable so that people do not try to open the door before the washing is over. The impact sound of a switch is always a good-feedback that a machine has been turned on.

Well-being. The physiology of the auditory system does not allow people to shut their ears. Consequently, people sometimes undergo product sounds involuntarily. Imagine computer fan sounds in offices, scooter sounds in the streets, or excessive alarm signals in intensive care units. Intensity of these sounds and their occurrence frequency may disturb people and can even cause fatigue.

User satisfaction. People often buy electric products without testing them in stores. They may be appealed by the product's appearance and functionality. However, when people try the product for the first time at home, they maybe disappointed by the unexpected sound. An electric toothbrush, for example, is a small object. People would not expect a roaring sound from it. To some user's disappointment some toothbrushes emit a roaring sound. Again, if the tray of a DVD player makes a rough friction sound, then users would be dissatisfied with their expensive DVD player because the sound does not match with their expectations.

Product identity. Congruency between a product and its sound is essential. Sound is expected to reflect the product characteristics. If a hairdryer looks very feminine, elegant, and powerful, people would expect the same qualities from its sound. A red sports car is expected to make a wilder sound compared to a serious-looking executive's car, which should probably function effortlessly and emit a sound on a comfortable level.

Brand differentiation. Products that belong to a single brand tend to have consistency in their appearance. Imagine how different Philips products look compared to Braun products. Sound, as well, is a property of a product, thus can be used as a means for brand differentiation. Consistent soundscape within the product range of a brand may even facilitate brand recognition in the visual absence of products.

Ultimately, it is the designers' task to fulfill the afore-mentioned needs to design sound. A well-designed product sound will be typical to the product and to the context, be informative about the product's operation cycle, and convey implicit/explicit characteristics of the product. In order to do so, design teams need to analyse the sound emitted by a product, determine the sound's function, conceptualize their ideas, and embody their decisions. They also need to understand how people would react to the sound in question and discuss about it in meetings.

Guidelines for sound design

How to start designing sounds

People's responses vary with respect to different product sounds (compare a shaver sound to an alarm clock sound). Accordingly, different categories of product sounds exist. These categories are indicative of concepts that characterize the product sounds. Thus, before designing sounds, determination of the category of the sound may be necessary. Assigning the category can be first done by the acoustical analysis of the sound because perceptually similar sounds tend to have similar physical disposition. Once the category is assigned, then basic concepts relevant to can be known. Designers can refer to Part A of this thesis to determine the category specific concepts. Once the these concepts are available, then designers can see beforehand how the sound in question will be responded to.

Communication of product sound characteristics

Designing as an activity requires in-depth communication about the characteristics of the product. Currently, product sound related vocabulary might be limited, probably because sound design is not a regular practice within the global design project. However, need for sound design is increasing, so is the need for sound related communication. The emerging 'sound' concepts provide examples of semantic associations that can be used to describe product sounds. However, the concepts do not provide a complete set of descriptive words. For example, *emotional responses* emerged as sound descriptions. This will require the understanding of the type of emotional response product sounds elicit. They could be similar to those emotions elicited by the appearance of a product and they could also be basic affective reactions to the auditory composition of the sound. How about sound-specific descriptions? When designing, communication of the exact auditory characteristics is essential to understand the physical composition of the sound. However, descriptions may not always refer to the auditory characteristics. For example, the word ‘round’ refers to a visual property or ‘hard’ refers to a tactile property of an object. That is to say, this thesis has provided the basics about sound descriptions. Still much is missing to understand the exact relations of these concepts to product sounds. The next step should be the determination of a sound-specific vocabulary that represents the occurring sound concept. Eventually, a lexicon specific to product sound descriptions will emerge.

Furthermore, recognition memory for product sounds can be hindered by the presence of a product. This may be a drawback for the verbal communication that elaborates on the acoustical properties of a sound. For better and more accurate communication sound designers need to capture the subtleties in the acoustic composition of the sound. Thus, listening to the sound in the absence of any product imagery will allow designers focus to on the sound. Consequently, the attention of the designers will not be split by any distracting images.

Visual images can be used to support the verbal communication of product sounds. However, their purpose should be to facilitate the retrieval of the product name. For example, during sound sketching and modeling, images of a product or its parts could be used as complimentary to the auditory experience. This would especially be an appropriate strategy for ambiguous sounds. Impact sounds, for example, refer to many events that have colliding parts. An image showing an action of impact (e.g., washing machine door closing) can disambiguate the label of the sound. Furthermore, pictograms are proposed to support the selection of the sounding product parts during sound modeling. It seems that designers will need higher perceptual expertise on the perceptual and semantic processing of the pictograms, which implies better training with pictograms and corresponding sounds.

Evaluation of product sounds

The fittingness of the sound to a product has often been assessed through abstract associations that derive from product characteristics (e.g., sportiveness, luxury, feminine). This thesis has shown that other types of meaningful associations exist that a product sound can be assessed with. Because sound’s relation to a product is not only based on abstract attributes, sound quality evaluations could be adapted to the purpose of the sound in a product. For example, some sounds are location-specific (e.g., shaver and bathrooms) whereas other sounds are used to provide feedback (e.g., mouse click). For some products, people want to distinguish between the events in a product (e.g., rinsing and drying cycles in dishwashers). These different sound functions provide the bases for sound quality evaluations.

Sound quality evaluations often take place in use-evaluation laboratories with no contextual cues offered. Because product sounds are inherently ambiguous, assessment of the sound without context may result in inappropriate semantic associations. However, identifying the sound correctly will provide associations directly about the product and its characteristics. For example, emotional responses can result from the sound as well as from the product itself. Thus, the pros-and-cons of the use of context during sound evaluation should be considered. Context can be provided by a photo of a location (e.g., bedroom) or by the presence of another object that is conceptually related to the product (hairbrush for a hairdryer).

Emotional responses

The same product sound can elicit positive or negative emotions depending on the identification stage from which these emotions result. For example, an espresso machine sound may be perceived as squeaky, too sharp and rough. In psychoacoustical terms, these sensations elicit negative affective responses. However, many people feel very comfortable and even happy in Italian cafés. The unpleasant sound from the espresso machine may be experienced as pleasant at a cognitive level. Thus, sometimes it is the sound that is typical to a product and even the sensory experience can be overcome by the concept of a product. Thus, design teams should consider the fittingness of a sound to a product or to the product experience.

One way to determine whether the emotional responses derive from the sound's acoustical composition or from the product itself would be to assess the sound and the product separately. This would allow designers to be aware of the potential causes of the positive or negative responses. A debriefing would be helpful to further determine the main cause of the responses. An additional evaluation of the sound together with the product could be necessary to determine the combined effect. Such systematic evaluation of emotions will also allow designers to specify which component of the product requires manipulation in order to achieve a desired sound.

Future suggestions

The knowledge that is presented in this thesis is based on empirical studies and supported by the theories of psychology. The intention is to support designers in their sound design related activities. Thus, a future study can be conducted to compare designers that make use of the knowledge presented in this thesis and other designers that design product sounds in their own intuitive way. Comparisons could concern efficiency, the quality of the design, and designers' (verbal) communication skills.

Simulation techniques such as sound sketching and sounding models have been proposed to support designers during the conceptual design phase. An important aspect is the evaluation of the result of simulation. With the knowledge on the concepts that a product sound may elicit, evaluation of the simulated sound will be easier and more realistic as these concepts directly reflect people's mental representation about product sounds.

Designing sounds should not be only for the person using the product. Other people passively undergoing sound should also be considered. Remember the scooter example. The rider on the scooter needs the sound for apparent reasons such as feedback. However, excessive hearing of such sound may cause discomfort. Thus, the psychological effect of sound on other people should also be studied to in order to develop methods that moderate the amount information that needs to be conveyed without compromising the well being of others.

These guidelines and suggestions are some indications given by the author. I am confident that designers and industry will reflect on this new knowledge and develop it further with their expertise in the field.

* * *

SUMMARY

Products are ubiquitous, so are the sounds emitted by products. Product sounds seem to influence our reasoning, emotional state, purchase decisions, preference, and expectations regarding the product and the product's performance. Thus, auditory experience elicited by product sounds may not be just about the act of hearing or a sensory response to an acoustical stimulus (e.g., this is a loud and sharp sound). People actually experience a product sound beyond its acoustical composition. People hear what the sound represents and appraise the product accordingly; or, they see what the product represents and appraise the sound accordingly.

Existing studies on product sounds mostly focused on the acoustic and engineering qualities of the sound in relation to the product and disregarded the human contribution to the experiential aspects of the sound. Determining the psychoacoustical reaction to a sound has been the next step engineers took to determine people's preference for certain sounds. In summary, our knowledge on product sounds is limited. A new approach, focusing on the psychological aspects of product sounds, is necessary to discover the meaning people derive from or attach to product sounds.

Understanding the human aspect of product sounds does not only concern the potential buyers or users. Ultimately, designers will benefit from this new approach. They will be able to predict the psychological consequences of their decisions and will be supported in their conceptual thinking regarding sounds. This new approach will provide a proper vocabulary that describes product sounds, and ultimately a systematic methodology to design sounds. Obviously, a gap exists between the

fundamentals of product sound experience and application of product sound design. This thesis bridges this gap by providing empirical findings and pointing out their relevance to the practice of product sound design.

Our knowledge about the world consists of concepts. In memory, these concepts consist of perceptual and semantic information concerning an object. Thus, seeing, hearing, feeling a product, interacting with it, or a being in a certain location will activate a bundle of relevant information that is glued by concepts. It seems impossible to isolate meanings attached to sounds from the influence of other product properties. Therefore, in this thesis product sounds are investigated through the concept of a product.

Similarly, early experiments (Chapters 1 and 2) investigate the concepts product sounds are represented with. First, basic sounds categories are determined based on the perceptual similarities of different product sounds. Accordingly, people can distinguish six categories of sounds: air, alarm, cyclic, liquid, impact, and mechanical. Each of these sound categories can be represented with various concepts in memory. Our studies suggest that eleven different types of basic concepts exist: action, emotion, location, material, abstract meanings, onomatopoeia, psychoacoustics, sound type, source, source properties, and temporal descriptions. These findings are the first to suggest that listeners' responses to product sounds are based on experiential aspects of sounds and not only on acoustical aspects. These experiential aspects also often relate to the product that emits the sound.

Considering the occurring product sound concepts, it is evident that it is often the product that dominates the mental representation of product sounds. To investigate this further, people's memory performances concerning product sounds are tested with accompanying pictures or text that described the product as sound source (Chapter 3). Interesting findings are that the presence of a picture or a label at learning a sound allows people remember the label of the sound. However, the presence of an image at learning a sound hinders the recognition performance. This suggests that source of the sound has a positive semantic influence on memory for product sounds, but negative influence on storing the sound's auditory properties (spectral-temporal).

As suggested by the first experiments, meaning attribution to product sounds occur on different levels of semantic association. The most commonly occurring sound description has been the 'source' description. That is, when people are asked to describe what they hear, their direct response would be labeling the sound by the product name. Despite their effort to label the product, people are not very good at providing the right label for the sound if the sound is presented without context

(Chapters 4 and 5). Incorrect responses often refer to a very similar product sound (e.g., hand-dryer instead of hairdryer). This makes product sounds ambiguous. However, the presence of context helps people to correctly label a sound (Chapter 5). The context may be a room in which the product sound often occurs (e.g., bedroom) or an object that is conceptually related to the product (e.g., hairbrush). The latter context provides the most information for the correct identification of the sound source.

In conclusion, these experimental findings altogether provide more insight into the mental representations of product sounds and demonstrate that responses given to product sounds depend on the type of the sound, availability of the context, and the use of interaction with the sound. Another important finding is that sound is an integral property of the product. Consequently, meaningful associations conveyed by a sound are subject to influences of the product concept. Furthermore, the ambiguity of the product sounds provides bases for the conceptual judgments. That is, a product sound may not be correctly identified as, e.g., a hairdryer. Yet, this sound will activate other concepts. Listeners use meanings derived from concepts to judge the congruency between a product and its sound.

The remaining part of the thesis mainly tackles designers' activities regarding product sounds. Because product sound design is a very new topic, (sound) designers lack tools and methods to design product sounds more efficiently. Therefore, a new visual tool that can facilitate the communication of sound designers during a design activity is proposed (Chapter 6). This tool makes use of pictograms that visually depict a composition of a product sound. Thus, a library of pictograms is designed to represent certain sound producing parts. With the sound producing parts and their physical representations on the computer, designers can model product sounds in the conceptual design phase.

In addition, the existing methods of product development are reviewed and a new methodology for designing product sounds is proposed (Chapter 7). Especially, sounding sketches and sounding models need to be included in a designer's daily routine of sound design. The proposed method enables (sound) designers to systematically tackle the sound design process and to efficiently communicate the sound related design problems/solutions.

Finally, the disciplines (acoustics, engineering, psychology, psychoacoustics, and musicology) contributing to product sound design are discussed and the responsibilities of a sound designer within the multi-disciplinary task of sound design are indicated (Chapter 8). It is suggested that product sound design should be an independent field that encompasses an inter-disciplinary approach. Therefore, design teams should include an expert who understands the inter-disciplinary nature of product sound design.

With this thesis, I hope to draw attention of both industry and academia to product sound design as an upcoming discipline. The thesis focuses on the human-aspect of product sounds. Findings demonstrate that the effects product sounds have on people are undeniable. Therefore, these effects should be considered at all times for the design, marketing, selling, and the use of the product.

Elif Özcan

SAMENVATTING

⁂

Producten zijn alomtegenwoordig en dat geldt ook voor het geluid dat zij produceren. Productgeluiden lijken invloed te hebben op ons logisch denken, onze emotionele toestand, ons koopgedrag, onze voorkeur en onze verwachtingen aangaande het product en zijn prestaties. Daarom is de auditieve ervaring van productgeluiden niet alleen het horen op zich, of een sensorische respons op een akoestische prikkel (bv. dit is een een hard of een scherp geluid). Het ervaren van een productgeluid overstijgt de akoestische samenstelling. Mensen geven betekenis aan een geluid en beoordelen het product dienovereenkomstig; ofwel zij zien een product en beoordelen het geluid dienovereenkomstig. Dat wil zeggen, er bestaat een complementaire en betekenisvolle verhouding tussen een product en zijn geluid.

Bestaande studies naar productgeluiden richtten zich meestal op de akoestische en technische eigenschappen van het geluid in relatie tot het product en veronachtzaamden de menselijke invloed op de ervaring van geluiden. Het meten van de psychoakoestische reactie op een geluid is voor ingenieurs bijvoorbeeld een beperkte extra stap in het bepalen van de voorkeur van mensen voor bepaalde geluiden. Samengevat: onze kennis over productgeluiden is beperkt. Er is dus een nieuwe aanpak nodig, één die zich richt op de psychologische aspecten van productgeluid. Deze aanpak is nodig om betekenissen die mensen ontlenen of toekennen aan productgeluiden te onderzoeken en te ontdekken.

Het begrijpen van de menselijke ervaring van productgeluiden staat niet alleen in relatie tot potentiële kopers of gebruikers. Uiteindelijk zullen vooral ook ontwerpers baat hebben bij deze nieuwe aanpak. Zij zullen mogelijk kunnen voorspellen wat de psychologische gevolgen van hun beslissingen aangaande productgeluiden zijn en zij worden ondersteund bij hun conceptuele denken over productgeluid. Vanuit dit onderzoek zal een vocabulaire ontwikkeld worden om productgeluiden te beschrijven en wordt ondersteuning geboden bij een systematische methodologie om geluiden te ontwerpen. Het is duidelijk dat er een zekere afstand zit tussen de beginselen van de ervaring van productgeluiden en de toepassing in het ontwerpen. Dit proefschrift tracht mede deze afstand te overbruggen door aan te geven welk belang de empirische resultaten voor de praktijk van het ontwerpen van productgeluiden kunnen hebben.

Onze kennis van de wereld bestaat uit representaties. In het geheugen wordt een relatie gelegd tussen de representatie en de perceptuele en semantische informatie van een object. Het zien, horen en voelen van een product, de interactie ermee, of het verblijven op een bepaalde plek, zal een verzameling relevante informatie activeren die gebundeld wordt in een representatie. Het is dus onmogelijk om betekenissen, die men aan het geluid toekent, los te zien van andere eigenschappen van het product. In dit proefschrift worden productgeluiden dan ook onderzocht via de representatie van een product.

Daarom zijn in de eerste experimenten (Hoofdstuk 1 en 2) de representaties van productgeluiden onderzocht. Eerst zijn primaire geluidscategorieën gedefinieerd op basis van de perceptuele gelijkenissen die verschillende productgeluiden kunnen hebben. Het blijkt dat mensen zes categorieën geluiden kunnen onderscheiden: *lucht, alarm, cyclisch, vloeibaar, impact,* en *mechanisch*. Elk van deze geluidscategorieën kan in het geheugen op verschillende wijze worden gerepresenteerd. Onze onderzoeken komen tot 11 verschillende types basis concepten: *actie, emotie, locatie, materiaal, abstracte betekenis, onomatopoësis (klanknabootsend), psychoakoestiek, geluidstype, bron, broneigenschappen,* en *temporele beschrijvingen*. Onze bevindingen geven aan dat de respons van luisteraars op productgeluiden in sterke mate is gebaseerd op ervaringsaspecten van geluiden en niet enkel op akoestische eigenschappen. Deze ervaringsaspecten hebben ook vaak betrekking op het product dat het geluid voort brengt.

Door de voorkomende representaties van productgeluid is het evident dat het vaak het product is dat de mentale representatie van productgeluiden domineert. Om dit verder te onderzoeken zijn de geheugenprestaties van mensen met betrekking tot productgeluiden getest met begeleidende afbeeldingen of met tekst die het product als geluidsbron beschreef (Hoofdstuk 3). Interessante resultaten zijn dat de aanwezigheid van een afbeelding of een begrip tijdens het leren mensen in staat stelt om het soort geluid te onthouden. Een dergelijke presentatie van een geluid bemoeilijkt echter de herkenning. Dit suggereert dat de bron van een geluid een positieve semantische invloed heeft op het geheugen voor productgeluiden, maar een negatieve invloed op het onthouden van de auditieve eigenschappen (spectraal-temporeel) van een geluid.

Zoals door de eerste experimenten is duidelijk geworden, vindt het toekennen van betekenis aan productgeluiden plaats op verschillende niveaus van semantische associaties. De meest voorkomende geluidsbeschrijving is de beschrijving van de ‘bron’. Dat wil zeggen dat mensen, indien gevraagd om te beschrijven wat ze horen, direct reageren met het benoemen van een geluid middels de productnaam. Ondanks hun inspanning om de relatie met een product te leggen,zijn mensen niet erg goed in het op juiste wijze benoemen van geluid als dat zonder context wordt aangeboden (Hoofdstuk 4 en 5). Onjuiste benoemingen vinden vaak plaats bij een zeer vergelijkbaar productgeluid (bijv. handendroger in plaats van föhn). Dit maakt het benoemen van productgeluiden vaag. Echter, de aanwezigheid van een context helpt mensen om een geluid correct te betitelen (Hoofdstuk 5). De context kan een kamer zijn waarin bepaalde productgeluiden vaak voorkomen (bijv. slaapkamer), of een object dat conceptueel aan het product verwant is (bijv. haarborstel). Deze laatste vorm van context blijkt de meeste informatie voor een juiste identificatie van de geluidsbron te geven.

Samengevat geven de resultaten van de experimenten meer inzicht in de mentale representaties van productgeluid en tonen zij aan dat reacties op productgeluid afhangen van: het type geluid , de beschikbaarheid van context, en de bedoeling achter de interactie met het geluid. Een andere belangrijk resultaat is dat geluid een integraal kenmerk is van het product. Betekenisvolle associaties die door het geluid worden overgebracht beinvloeden het productconcept. Bovendien verschaft de ambiguïteit van productgeluiden in zekere zin de basis voor een conceptueel oordeel. Dat wil zeggen dat een productgeluid wellicht niet op correct wijze wordt geïdentificeerd als, bijvoorbeeld, een föhn. Toch zal dit geluid andere representaties activeren. Met de betekenis die van deze representaties is afgeleid zullen luisteraars beslist de overeenstemming tussen een product en zijn geluid beoordelen.

Het resterende deel van het proefschrift heeft betrekking op de relatie tussen product ontwerpen en productgeluid. Omdat het ontwerpen van productgeluid een zeer nieuw onderwerp is, missen (geluids)ontwerpers vooralsnog het gereedschap en de methodes om productgeluid efficiënt te ontwikkelen. Daarom is een visueel hulpmiddel ontwikkeld, dat de communicatie met geluidsontwerpers kan ondersteunen (Hoofdstuk 6). Dit gereedschap maakt gebruik van pictogrammen die een samenstelling van een productgeluid voorstellen. Zo is een set pictogrammen ontworpen voor bepaalde onderdelen. Met deze geluidsproducerende onderdelen en hun fysieke representaties op de computer kunnen ontwerpers productgeluiden modelleren.

Aansluitend is er gekeken naar bestaande methodes voor productontwikkeling en is een nieuwe methodologie voor het ontwerpen van productgeluiden voorgesteld (Hoofdstuk 7). Vooral het hanteren van geluidsschetsen en geluidsmodellen dienen opgenomen te worden in de dagelijkse praktijk van het geluidsontwerpen. De voorgestelde methode verschaft (geluids)ontwerpers de mogelijkheid om systematisch het proces van geluidsontwerp aan te pakken en om efficiënt te communiceren over geluidsgerelateerde problemen/oplossingen.

Tenslotte zijn de disciplines die bijdragen aan het ontwerpen van productgeluiden (akoestiek, constructie, psychologie, psychoakoestiek en musicologie) geanalyseerd en is de verantwoordelijkheid van een geluidsontwerper binnen de multidisciplinaire taak van het geluidsontwerp omschreven (Hoofdstuk 8). Voorgesteld wordt dat het ontwerpen van productgeluid beschouwd dient te worden als een zelfstandig gebied dat om een interdisciplinaire aanpak vraagt. Daarom dienen ontwerpteams een expert te bevatten die het interdisciplinaire karakter van het ontwerpen van productgeluiden begrijpt.

Met dit proefschrift hoop ik de aandacht van zowel de industrie als de academische wereld te richten op het ontwerpen van productgeluid als een discipline in opkomst. Dit proefschrift richt zich op het menselijke aspect van productgeluiden. De bevindingen tonen aan dat de psychologische effecten van productgeluiden op mensen onmiskenbaar zijn. Daarom dienen deze effecten ten allen tijde in beschouwing te worden genomen bij het ontwerpen, in de markt zetten, verkopen, en gebruiken van het product.

Elif Özcan

SOCIETAL IMPACT

⁂

NRC HANDELSBLAD – 2008

Elif Ozcan Vieira, ontwerper van productgeluiden

'Geluid is geen lawaai'

Ontwerpers besteden te weinig aandacht aan het geluid dat een product maakt. Dat vindt ir. Elif Ozcan Vieira (Industrieel Ontwerpen). Ze deed onderzoek naar de beleving van productgeluiden en ontwikkelde een methode om geluid van huishoudelijke apparaten te ontwerpen.

MAAIKE MULLER

INTERVIEW

Foto: Sam Rentmeester/FMAX

Naast haar studie industrieel ontwerpen in Turkije, maakte Elif Ozcan Vieira jingles en geluidseffecten bij een radiostation. "Ik kwam erachter dat geluid ontwerpen helemaal niet zo anders is dan visueel ontwerpen." Na een aantal jaren als multimedia-ontwerper in Lissabon, stortte ze zich in Delft op het relatief nieuwe onderzoeksgebied van geluidsontwerpen. Voor haar promotieonderzoek ontwikkelde ze een methode om het geluid van een product te ontwerpen.

Waarom moet een ontwerper het geluid van huishoudelijke apparaten ontwerpen?

"Het geluid dat een product maakt, als je het gebruikt of bijvoorbeeld ergens neerzet, moet passen bij dat product. Het moet een positieve ervaring oproepen. Gebruikers vinden dat belangrijk, in een tijd dat het met de functionaliteit van het product wel goed zit. Een koffieapparaat moet plezierig klinken, want mensen associëren een kopje koffie drinken met ontspanning."

Houden ontwerpers geen rekening met geluid?

"Veel te weinig. Ze houden zich een beetje bezig met geluiden, zoals het piepen van een wekker. Maar geluiden die horen bij het gebruik van bijvoorbeeld een stofzuiger, de klik van het knopje als je hem aanzet of het geluid van de motor, krijgen weinig aandacht. De enige uitzondering is de auto-industrie, waar veel energie en geld wordt gestoken in het ontwerpen van het juiste geluid van de motor of het dichtslaan van deuren. Ik vind dat ontwerpers zich ervan bewust moeten zijn dat het geluid van een product verandert als het ontwerp, bijvoorbeeld het materiaal, verandert."

Dat klinkt logisch, maar hoe ontwerp je geluid?

"Tijdens mijn promotie heb ik een methode ontwikkeld voor het ontwerpen van productgeluiden. Het proces van geluid ontwerpen gaat in dezelfde stappen en parallel aan het ontwerpproces voor het product. Twee stappen in de methode zijn echt nieuw voor geluidsontwerpen. Een daarvan is het maken van geluidsschetsen. Hiervoor verzamelt de ontwerper objecten die geluid maken. Hij analyseert hoe goed die passen bij het product. Het komt overeen met de visuele schetsen die een ontwerper in het begin van een proces maakt om de vorm te bepalen. Daarna maakt een ontwerper vaak een 3D-model in de computer om het ontwerp beter te zien. Analoog daaraan maakt hij een model van het geluid. We hebben software ontwikkeld, waarmee de ontwerper in een soort bibliotheek geluiden kan kiezen voor alle geluidmakende onderdelen van een product. De motor, een klep die dicht gaat of een knopje. Met de computer kan het geluid van het hele product worden gesimuleerd, de *soundscape*."

Hoe wordt je methode ontvangen?

"Studenten die ik ermee heb laten werken, vinden het leuk en nuttig. En ook in de academische wereld en de industrie is veel interesse. Grote productontwerpende bedrijven wisten ze niet hoe ze geluid moesten ontwerpen, dus lieten ze het maar zitten. Over kleur en vorm bestaan misschien al eeuwen theorieën, maar over geluid niet. Ik ben een van de weinigen die onderzoek doet naar geluidsontwerpen. Er moet nog veel kennis worden ontwikkeld, bijvoorbeeld over de emotionele respons op geluid van een product. Geluid moet bij het product passen. Het is heel duidelijk niet mijn bedoeling geluid als lawaai te zien en het product zo te ontwerpen dat het zo min mogelijk geluid maakt. Mijn onderzoeksgroep gaat bijvoorbeeld meedenken over het geluid van een auto op waterstof. Auto's kunnen heel stil zijn. Dat lijkt fijn, maar je wilt wel graag horen of de motor aan staat en wanneer je sneller gaat."

Meer informatie: E. Ozcan Vieira, e.ozcan@tudelft.nl.

Geen design zonder geluid

Sound design is een snel opkomend vak. En niet alleen de auto-industrie dingt met sound naar de gunst van de consument. Hoe gaat dat, geluid ontwerpen?

Door onze redacteur
WARNA OOSTERBAAN

MÜNCHEN, 13 AUG. „Luistert u hier eens even naar.” Dr. Gerhard Thoma is psycho-akoesticus. Hij drukt op een knop. In een klein auditorium van het *Forschungs und Innovations Zentrum* van de Duitse autofabrikant BMW in München klinkt het geluid van een autoraam dat zich sluit.

„Hoorde u hoe de toonhoogte halverwege wat zakte, en aan het einde weer omhoogging?”

Ja, dat hoorden we.

„Dat is heel verklaarbaar”, vervolgt Thoma, een vijftiger met een professorale voordracht. „De mensen willen graag autoruiten die aan de bovenkant een beetje bol staan, dat ziet er mooier uit. Het gevolg is wel dat de ruit halverwege een beetje klem in de sponning komt te zitten en er pas aan het eind weer mooi inpast. De elektromotor voor de raambediening moet dus in het midden harder werken, en zakt iets in toerental en in toonhoogte. Constructief gezien is dat geen enkel probleem. Maar de klant krijgt de indruk dat die motor het maar nét trekt, en dat is strijdig met het beeld dat ze van een BMW hebben.”

Wat deed BMW aan dit probleem? Ze verbeterden de sponning en ze maakten de elektromotor zo sterk dat die de ruit zonder aarzeling door de sponning sleurt.

Waarom? Omdat, zo legt Thoma uit, geluid steeds belangrijker wordt. „Na het uiterlijk van een auto is het geluid het eerste dat een aspirant-koper opvalt. De betrouwbaarheid, de veiligheid en de milieuvriendelijkheid zijn ook allemaal belangrijke factoren, maar een nieuwe klant die een proefrit maakt, merkt daar niets van. Je moet hem pakken met het uiterlijk, met *styling* en met geluid.”

Daarom doet BMW – al sinds 1991 – aan *sound design*. In het uitgestrekte onderzoeks- en innovatiecentrum van de autofabrikant, waar negenduizend ingenieurs en technici aan de auto's 'van morgen' werken, geeft Thoma leiding aan honderdtwintig geluidsspecialisten. „Wij ontwerpen geluiden die niet authentiek zijn, maar die de klanten voor authentiek houden”, legt hij uit.

Daarom worden de ruitenwissermotoren zo gemaakt dat ze niet in toonhoogte op en neer gaan – hoewel het mechanisme, dat met een heen en weer bewegend excentriek werkt, daar wel alle aanleiding toe geeft. En daarom geeft het dashboardkastje van de Duitse autofabrikant een vertrouwenwekkende klak.

Sound design is een snel opkomend vak. En niet alleen in de auto-industrie. Ook in andere bedrijfstakken wordt steeds minder met puur technische eigenschappen van het product naar de gunst van de consumenten gedongen. De *look*, de *feel* en zeker ook de *sound* – daar komt het op aan.

Vervolg **Geluid**: pagina 14

NRC HANDELSBLAD (COVER STORY) – 2008

Een te stille stofzuiger verkoopt niet

Vervolg **Geluid** van pagina 1

Thoma van BMW geeft een ander voorbeeld: „Een autodeur is van dun plaatstaal gemaakt, hij is maar 0,7 mm dik. Heel normaal dat je een blikkerig geluid krijgt als je die dichtslaat. Maar een BMW-rijder verwacht een solide klik. Dus wij monteren antidreun-matten, en we passen de constructie van het portierslot aan.”

Aan de TU Delft is industrieel ontwerper Elif Özcan (32) bezig sound design van een academisch fundament te voorzien. Over twee maanden promoveert zij. Ze deed onderzoek naar de manier waarop productgeluiden zich in de menselijke geest manifesteren, en probeert met die inzichten ontwerpers verder te helpen. „Aanvankelijk was geluid vooral iets dat bestreden moest worden, bijvoorbeeld bij stofzuigers. Ze maken nu veel minder geluid dan vroeger.” Maar daar kun je te ver mee gaan, ontdekten de stofzuigerfabrikanten. „Als een stofzuiger heel stil is, denken de mensen dat die niet werkt.”

Eigenlijk moet je bij elk ontwerp vanaf het begin aandacht besteden aan de rol van het geluid, vindt Özcan. Want het geluid is voor een gebruiker niet iets wat erbij komt, maar is deel van de identiteit van dat product. „Hoe een apparaat eruitziet is het eerste beslissende moment, maar het tweede komt als je de stekker in het stopcontact steekt en hoort welk geluid het maakt. Het team dat het Senseo-koffiezetapparaat heeft ontworpen kwam er pas laat achter dat het apparaat niet goed klonk, te goedkoop. Ze hebben er net zo lang aan gesleuteld tot het wel goed klonk.”

Door het geluid te ontwerpen kunnen ontwerpers een emotionele lading aan een product geven, zegt Özcan, „denk maar de klappen die een Harley Davidson-motorfiets maakt, het is een belangrijk deel van de identiteit van dat merk.

Maar ze kunnen met geluid ook informatie overdragen – een bliepje dat je vertelt dat je een toets van je telefoon goed hebt ingedrukt, een reeks piepjes als je de veiligheidsgordels niet omdoet, of met het opgewekte plieng als de magnetron klaar is met je eten. Dat moet allemaal ontworpen worden.”

Maar hoe doe je dat, geluid ontwerpen? Hoe praat je erover? Özcan onderzocht welke soorten geluiden mensen onderscheiden, ze ontwierp pictogrammen voor elke categorie, en ze werkt aan een „bibliotheek van geluiden”, opgenomen geluiden van allerlei industriële onderdelen, met behulp waarvan een ontwerper een geluid kan simuleren.

Bij BMW wordt bij het ontwerp van de automotoren al in die richting gewerkt. Thoma: „Een luxe personenwagen moet zijdezacht lopen, een sportwagen mag best wat rauwer klinken en bij een SUV

luiden nu op de computer simuleren en op die manier beslissen hoe een nieuwe auto moet gaan klinken.”

De motortechnici bij BMW krijgen bij de ontwikkeling van nieuwe modellen de geluidssimulaties nu als huiswerk mee, en het is hun taak aan het gewenste profiel te voldoen. Hun belangrijkste instrumenten zijn de vorm van de luchtinlaat – de *airbox* – en de vorm en ligging van de uitlaat. Thoma: „Het is alleen wel jammer dat hoe zuiniger je een motor maakt, des te rauwer hij gaat klinken. Daar is niet zoveel aan te doen.”

Ook bij Philips is geluid aan een opmars bezig – hoewel sound design er geen aparte discipline is. Maar sinds het bedrijf de richting van '*sense and simplicity*' is ingeslagen en bezig is zich te ontwikkelen tot een bedrijf dat zich minder op techniek en meer op consumenten richt, sinds die tijd speelt ook het geluid dat al die apparaten maken een steeds belangrijker rol, zegt Thomas Marzano, *creative director* bij Philips Design in Eindhoven. Bij medische apparatuur als CRT-scanners wordt er alles aan gedaan de meest angstaanjagende geluiden te dempen, bij de nieuwe breedbeeldtelevisies wordt het geluid ook zo breed mogelijk gemaakt, en waar mogelijk worden scheerapparaten en citruspersen uitgerust met aangenaam gebrom in plaats van gesnerp en gegier.

> 'De Senseo klonk eerst te goedkoop'
> **Elif Özcan, TU Delft**

„We zijn er al lang mee bezig. We hebben al heel vroeg de cd-laatjes zo gemaakt dat ze soepel klinken, en geen plastic geratel produceren ”, zegt Marzano.

Maar je kunt er ook te ver in gaan. „Toen de Philips Whirlpool wasmachines werden uitgerust met antigeluid – een voorziening die door het produceren van geluid in tegenfase hinderlijk geluid kan onderdrukken – kregen we veel wasmachines terug met kapot gerukte handels. De machines waren zo stil geworden dat de mensen dachten dat ze het niet deden.”

Elif Özcan van de TU Delft heeft ook ervaringen met een stil apparaat: „Ik rijd in een Toyota Prius, en als je op een parkeerterrein manoeuvreert horen de mensen je niet, want dat doet hij op zijn elektromotor. Ik heb al een paar keer gehad dat voetgangers schrokken omdat ze mijn auto niet hadden opgemerkt. Blijkbaar houd je er onbewust rekening mee dat je oren je wel zullen waarschuwen.”

De huidige verkoophit van Philips, de Wake-up Light bracht een ander akoestisch probleem. Bij deze wekker, die wekt met licht en met een geluid dat zelf is in te stellen, wordt door veel kopers het gefluit van vogels als het favoriete wekgeluid gekozen. Dat werkt

Potatoe-potatoe-potatoe

Behalve met het júiste geluid willen fabrikanten zich ook met een éigen geluid onderscheiden. Zoals bierfabrikant Grolsch met de bekende plop van de beugelfles. De opstart- en afsluitgeluiden van de verschillende versies van Windows behoren tot de beroemdste sound logo's van de afgelopen decennia.

Sommige fabrikanten patenteren hun geluid. Filmproducent MGM heeft patent op zijn leeuwenbrul, chipfabrikant Intel op zijn bekende riedeltje, Nokia op zijn standaard ringtone.

De Amerikaanse motorfabrikant Harley Davidson heeft in 1994 geprobeerd de Japanse concurrentie een slag toe te brengen door de syncopische roffel van zijn V-twin motoren (volgens Harley-fans klinkt die als potatoe-potatoe-potatoe) te patenteren, maar is daarin niet geslaagd. Het geluid bleek niet specifiek genoeg.

Bij Philips is hard gestudeerd op een herkenningsmelodie om de slogan *sense and simplicity* te onderstrepen. Marketingman Gary Raucher, verantwoordelijk voor het sound logo: „We wilden een simpel geluid, van één instrument en in twee tonen, om die twee woorden te versterken. We hebben honderden combinaties onderzocht.” De uiteindelijke versie – twee triangelachtige tonen – wordt nu in alle Philips-*commercials* gebruikt. In het Braziliaanse kantoor van Philips weerklinkt zij wanneer een lift halt houdt bij de gewenste verdieping.

Zeven basisgeluiden

Hiernaast pictogrammen voor productgeluiden die industrieel ontwerper Elif Özcan van de TU Delft ontwikkelde om de communicatie over het ontwerpen van geluiden te vergemakkelijken. Zij onderscheidt zeven basisgeluiden: gepiep, het geluid van vloeistof (druppelen, stromen), het zoemende geluid van een motor, ander mechanische geluiden (transmissie, tandwielen), contactgeluiden (stoten, bonken), het geluid van een ventilator en geklik. Door variaties in de afbeelding van de pictogrammen (bij voorbeeld donkerder of lichter) kunnen variaties in de geluiden worden verbeeld, zoals verschillen in

NRC HANDELSBLAD – 2008

Misschien toch beter om te reanimeren

▶ Opinie: pagina 16

Eindelijk weer! De oerversie van Stravinsky

▶ Muziek: pagina 28

nrc▸tv

nrc kamatube

Parkman

nrc•next

woensdag 20 augustus 2008
www.nrcnext.nl
prijs 1 euro

Uitgave van NRC Handelsblad Jaargang 3 No.110

Zo klinkt een Harley

Het geluid van de wasmachine, de stofzuiger en de autodeur klinkt authentiek, maar is dat niet. Grote bedrijven hebben specialisten in dienst die dat geluid ontwerpen.

▶ pagina 4 en 5

Illustratie Robin Héman

Nederland
Reservisten met oud materieel

Het Korps Nationale Reserve kampt met verouderd materieel. Er zijn vooral problemen met bewapening, nachtzichtapparatuur en communicatiemiddelen. Een aantal reservisten moest lichte machinegeweren inruilen voor gewone geweren. De machinegeweren gingen naar Uruzgan.

▶ pagina 3

Internationaal
NAVO hard over actie Rusland

De NAVO is het gisteren eens geworden over een veroordeling van Rusland. De NAVO-landen willen voorlopig geen contact met Rusland in de zogenoemde 'NAVO-Ruslandraad'. Dat gebeurde vooral onder druk van de Verenigde Staten, Groot-Brittannië en de Oost-Europese landen.

▶ pagina 6 en 7

Economie
Dure hypotheek voor oversluiter

Enkele honderdduizenden huizenbezitters in Nederland zijn straks honderden euro's per maand extra kwijt aan hun nieuwe, 'overgesloten' hypotheek. Indertijd sloten zij een hypotheek af met een looptijd van vijf jaar tegen een relatief lage rente. Die rente is nu fors gestegen.

▶ pagina 10 en 11

Sport
Van Grunsven weer de beste

Amazone Anky van Grunsven heeft voor de derde keer olympisch goud gewonnen bij de dressuur. Ze won ook goud in 2000 en 2004. De waterpolovrouwen spelen morgen de finale na een zege tegen Hongarije, en turner Epke Zonderland werd gisteren zevende aan rek.

▶ pagina 12 t/m 15, 18 en 19

nrcnext.nl/abo
4 weken vanaf 9,-
Be next

33408
8 710114 600035

NRC.NEXT (COVER STORY) – 2008

Geluid dat vertrouwen wekt

Met sounddesign zorgen fabrikanten ervoor dat alle apparaten goed klinken

▸ Van een apparaat dat geen geluid maakt, denken mensen dat het kapot is.

▸ Dus voegen fabrikanten geluiden toe: een solide klik, een geruststellend gebrom.

Door WARNA OOSTERBAAN

MÜNCHEN. „Luistert u hier eens even naar." Dr. Gerhard Thoma is psycho-akoesticus. Hij drukt op een knop. In een klein auditorium van het *Forschungs und Innovations Zentrum* van de Duitse autofabrikant BMW in München klinkt het geluid van een autoraam dat zich sluit.

„Hoorde u hoe de toonhoogte halverwege wat zakte, en aan het einde weer omhoogging?"

Ja, dat hoorden we.

„Dat is heel verklaarbaar", vervolgt Thoma. „De mensen willen graag autoruiten die aan de bovenkant een beetje bol staan, dat ziet er mooier uit. Het gevolg is wel dat de ruit halverwege een beetje klem in de sponning komt te zitten en er pas aan het eind weer mooi inpast. De elektromotor voor de raambediening moet dus in het midden harder werken, en zakt iets in toerental en in toonhoogte. Constructief gezien is dat geen enkel probleem. Maar de klant krijgt de indruk dat die motor het maar nét trekt – en dat is strijdig met het beeld dat ze van een BMW hebben."

Wat deed BMW aan dit probleem? Ze verbeterden de sponning en ze maakten de elektromotor zo sterk dat die de ruit zonder aarzeling door de sponning sleurt.

Waarom? Omdat, legt Thoma uit, geluid steeds belangrijker wordt. „Na het uiterlijk van een auto is het geluid het eerste dat een aspirant-koper opvalt. De betrouwbaarheid, de veiligheid en de milieuvriendelijkheid zijn ook allemaal belangrijke factoren, maar een nieuwe klant die een proefrit maakt, merkt daar niets van. Je moet hem pakken met het uiterlijk, met *styling* en met geluid."

Daarom doet BMW aan *sound design* – al sinds 1991. In het uitgestrekte onderzoeks- en innovatiecentrum van de autofabrikant, waar negenduizend ingenieurs en technici aan de auto's 'van morgen' werken, geeft Thoma leiding aan honderdtwintig geluidsspecialisten. „Wij ontwerpen geluiden die niet authentiek zijn, maar die de klanten voor authentiek houden", legt hij uit.

Daarom worden de ruitenwissermotoren zo gemaakt dat ze niet in toonhoogte op en neer gaan – hoewel het mechanisme, dat met een heen en weer bewegend excentriek werkt, daar wel alle aanleiding toe geeft. En daarom geeft het dashboardkastje een vertrouwenwekkende klak.

Sound design is een snel opkomend vak, niet alleen in de auto-industrie. Ook in andere bedrijfstakken wordt steeds minder met puur technische eigenschappen van het product naar de gunst van de consumenten gedongen. De *look*, de *feel* en zeker ook de *sound* – daar komt het op aan.

Thoma van BMW geeft een ander voorbeeld: „Een autodeur is van dun plaatstaal gemaakt, hij is maar 0,7 mm dik. Heel normaal dat je een blikkerig geluid krijgt als je die dichtslaat. Maar een BMW-rijder verwacht een solide klik. Dus wij monteren antidreunmatten en we passen de constructie van het portierslot aan."

Aan de TU Delft is industrieel ontwerper Elif Özcan (32) bezig sound design van een academisch fundament te voorzien. Over twee maanden promoveert ze. Özcan deed onderzoek naar de manier waarop productgeluiden zich in de menselijke geest manifesteren en probeert met die inzichten ontwerpers verder te helpen. „Aanvankelijk was geluid vooral iets wat bestreden moest worden, bijvoorbeeld bij stofzuigers. Ze maken nu veel minder geluid dan vroeger." Maar daar kun je te ver mee gaan, ontdekten de stofzuigerfabrikanten. „Als een stofzuiger heel stil is, denken mensen dat-ie niet werkt."

▸ De Senseo klonk eerst te goedkoop. Daar is lang aan gesleuteld ◂

Elif Özcan, geluidsonderzoeker

Eigenlijk moet je bij elk ontwerp vanaf het begin aandacht besteden aan de rol van het geluid, vindt Özcan. Want het geluid is voor een gebruiker niet iets wat erbij komt, maar is deel van de identiteit van dat product. „Hoe een apparaat eruitziet, is het eerste beslissende moment, maar het tweede komt als je de stekker in het stopcontact steekt en hoort welk geluid het maakt. Het team dat het Senseo-koffiezetapparaat heeft ontworpen, kwam er pas laat achter dat het apparaat niet goed klonk: te goedkoop. Ze hebben er net zo lang aan gesleuteld tot het wel goed klonk."

Door het geluid te ontwerpen kunnen ontwerpers een emotionele lading aan een product geven, zegt Özcan. „Denk maar aan de klappen die een Harley Davidson-motorfiets maakt, het is een belangrijk deel van de identiteit van dat merk. Maar ze kunnen met geluid ook informatie overdragen – een bliepje dat je vertelt dat je een toets van je telefoon goed hebt ingedrukt, een reeks piepjes als je de veiligheidsgordels niet omdoet, of het opgewekte *plieng* als de magnetron klaar is met je eten. Dat moet allemaal ontworpen worden."

Maar hoe doe je dat, geluid ontwerpen? Hoe praat je erover? Özcan onderzocht welke soorten geluiden mensen onderscheiden, ze ontwierp pictogrammen voor elke categorie en ze werkt aan een „bibliotheek van geluiden", opgenomen geluiden van allerlei industriële onderdelen, met behulp waarvan een ontwerper een geluid kan simuleren.

Bij BMW wordt bij het ontwerp van de automotoren al in die richting gewerkt. Thoma: „Een luxe personenwagen moet zijdezacht lopen, een sportwagen mag best wat rauwer klinken en bij een SUV moet je de indruk krijgen dat die zich door elke modderpoel heen kan werken. We kunnen al die geluiden nu op de computer simuleren en op die manier beslissen hoe een nieuwe auto moet gaan klinken."

De motortechnici bij BMW krijgen bij de ontwikkeling van nieuwe modellen de geluidssimulaties nu als huiswerk mee en het is hun taak aan het gewenste profiel te voldoen. Hun belangrijkste instrumenten zijn de vorm van de luchtinlaat – de *airbox* – en de vorm en ligging van de uitlaat. Thoma: „Het is alleen wel jammer dat hoe zuiniger je een motor maakt, des te rauwer hij gaat klinken. Daar is niet zo veel aan te doen."

Ook bij Philips is geluid aan een opmars bezig – hoewel sound design er geen aparte discipline is. Maar sinds het bedrijf de richting van '*sense and simplicity*' is ingeslagen en bezig is zich te ontwikkelen tot een bedrijf dat zich minder op techniek en meer op consumenten richt, sinds die tijd speelt ook het geluid dat al die apparaten maken een steeds belangrijker rol, zegt Thomas Marzano, *creative director* bij Philips Design in Eindhoven. Bij medische apparatuur als CRT-scanners wordt er alles aan gedaan de meest angstaanjagende geluiden te dempen, bij de nieuwe breedbeeldtelevisies wordt het geluid ook zo breed mogelijk gemaakt, en waar mogelijk worden scheerapparaten en citruspersen uitgerust met aangenaam gebrom in plaats van gesnerp en gegier.

„We zijn er al lang mee bezig. We hebben al heel vroeg de cd-laatjes zo gemaakt dat ze soepel klinken en geen plastic geratel produceren", zegt Marzano. Maar je kunt er ook te ver in gaan. „Toen de Philips Whirlpool wasmachines werden uitgerust met antigeluid – een voorziening die het hinderlijke geluid opheft door een tegengestelde geluidsgolf te produceren – kregen we veel wasmachines terug met kapot gerukte handels. Ze waren zo stil geworden dat mensen dachten dat ze het niet deden."

Elif Özcan heeft ook ervaringen met een stil apparaat: „Ik rijd in een Toyota Prius en als je op een parkeerterrein manoeuvreert, horen de mensen je niet, want dat doet hij op zijn elektromotor. Ik heb al een paar keer gehad dat voetgangers schrokken, omdat ze mijn auto niet hadden opgemerkt. Blijkbaar houd je er onbewust rekening mee dat je oren je wel zullen waarschuwen."

De huidige verkoophit van Philips, de Wake-up Light, bracht een ander akoestisch probleem. Bij deze wekker, die wekt met licht en met een geluid dat zelf is in te stellen, wordt door veel kopers het gefluit van vogels als het favoriete wekgeluid gekozen. Dat werkt heel goed, tot het lente wordt en de mensen elke dag om half vijf wakker worden – van echte vogels.

Potato-potato-potato

Behalve met het *juiste* geluid willen fabrikanten zich ook met een *eigen* geluid onderscheiden. Zoals bierfabrikant Grolsch dat bijvoorbeeld doet met de bekende plop van de beugelfles. De opstart- en afsluitgeluiden van de verschillende versies van Windows behoren tot de beroemdste 'sound logo's' van de afgelopen decennia.

▶ **Lees** over Bill Brown, componist van geluidjes voor Windows XP, op nrcnext.nl/links

Sommige fabrikanten patenteren hun geluid. Filmproducent MGM heeft patent op zijn leeuwenbrul, chipfabrikant Intel op zijn bekende riedeltje, Nokia op zijn standaard ringtone.

▶ **Beluister** de MGM-leeuw en lees over de oorsprong van de Nokia-ringtone op nrcnext.nl/links

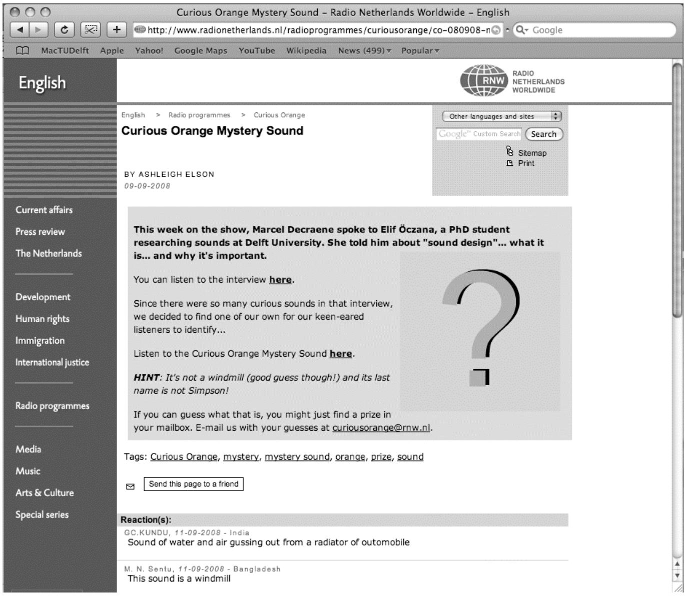
De Amerikaanse motorfabrikant Harley Davidson heeft in 1994 geprobeerd de Japanse concurrentie een slag toe te brengen door de syncopische roffel van zijn V-twin motoren (volgens Harley-fans klinkt die als potato-potato-potato) te patenteren, maar is daarin niet geslaagd. Het geluid bleek niet specifiek genoeg.

▶ **Bekijk** een filmpje met het geluid via nrcnext.nl/links.

Bij Philips is hard gestudeerd op een herkenningsmelodie om de slogan *sense and simplicity* te onderstrepen. Marketingman Gary Raucher, verantwoordelijk voor het sound logo: „We wilden een simpel geluid, van één instrument en in twee tonen, om die twee woorden te versterken. We hebben honderden combinaties onderzocht." De uiteindelijke versie – twee triangelachtige tonen – wordt nu in alle Philips-*commercials* gebruikt. In het Braziliaanse kantoor van Philips weerklinkt zij wanneer een lift halt houdt bij de gewenste verdieping.

Illustratie Robin Héman en Ingrid van Halteren

NRC.NEXT – 2008



English > Radio programmes > Curious Orange

Curious Orange Mystery Sound

BY ASHLEIGH ELSON

09-09-2008

This week on the show, Marcel Decraene spoke to Elif Öczana, a PhD student researching sounds at Delft University. She told him about "sound design"... what it is... and why it's important.

You can listen to the interview **here**.

Since there were so many curious sounds in that interview, we decided to find one of our own for our keen-eared listeners to identify...

Listen to the Curious Orange Mystery Sound **here**.

***HINT**: It's not a windmill (good guess though!) and its last name is not Simpson!*

If you can guess what that is, you might just find a prize in your mailbox. E-mail us with your guesses at curiousorange@rnw.nl.

Tags: Curious Orange, mystery, mystery sound, orange, prize, sound

Reaction(s):

GC.KUNDU, *11-09-2008* - India
Sound of water and air gussing out from a radiator of outomobile

M. N. Sentu, *11-09-2008* - Bangladesh
This sound is a windmill

Radio Netherlands (www.rnw.nl) – 2008

ACKNOWLEDGMENTS

"All the world's a stage
And all the men and women merely players;
They have their exits and their entrances,
And one man in his time plays many parts,
His acts being seven ages."
William Shakespeare, As You Like It, 2/7

This is the end of the third act in my ongoing play. The first act was about me – me discovering who I am and me shaping up. The second act was about him – him loving me and taking me away. The third one was about *it* – it taking over me and him, it becoming a passion, it becoming a compulsion, it broadening my view on life, it satisfying my intellectual needs, and it making me feel useful. It was a dream and it came true.

I have never been alone in my own play. In the first and second acts my parents were with me. They taught me how to be strong. They exemplified why I should go after my passions. They never asked questions except the one 'if that is going to make you happy'. In the second act, I was mostly with Alex. He showed me how it feels like to love and be loved in return. He taught me that life without risks is not fun. He challenged me with his ideas and comforted me with his kisses and hugs. In the third act I was bound to be alone by the nature of the project. It was my thing and I had to do it myself. But I wasn't. I had my parents in my thoughts, my love next to me, and many others circling me, everybody supporting me from all directions.

I met the main players of the third act in my job interview: Jan Jacobs, my promotor; René van Egmond, my supervisor, and Paul Hekkert, my current boss. The first impressions I had that day still rule all my decisions about these gentlemen. There was a great energy in the room. Talking to them was like talking to a good friend or to a family member. I went back home completely bewildered and praying that I'd get the job. Then I thought even if I didn't at least I had great time in there. I got the job and these three wise men have become the most influential. Jan was definitely the father figure in the whole act. I felt so protected with him and taken care of. René was the *super* supervisor. He always had a smile and a big 'hi'. After my mother, he has become the second voice in my head. Paul has become a friend, not a boss - well, let's put it this way: a friendly boss. These men trusted me all the way and helped me build my career. I owe a lot to them.

There were also other big players that helped my academic development. Huib de Ridder, Ans Koenderink, Norbert Roozenburg, Petra Badke-Schaub, and Rick Schifferstein have been great teachers about scientific thinking. Judy Edworthy has provided me with a place at the Plymouth University and showed me how psychologists think. James Ballas gave me his time and attention.

I got technical support from various people. I often thought Rob, the Rijswijker, lost all his hair because he worked on my stubborn button-box. Aadjan gave his technical support with nice chats on the side. Corrie has been great help with the preparation of this book for publishing. Carlita and Timo were always a good sport!

I spent countless hours at my desk behind my computer. How lucky I was to have great roomies such as Jenneke, Jeroen, and Marieke. Jenneke has become my daily addiction and my common sense. We laughed together as much as we cried together. Jeroen has become a great friend. I had the most interesting discussions with him both about science and about the dynamics of men-women relationship. Marieke's never-ending energy filled me up with energy. Her compassionate support gave me hope that I could finish *it*.

Paul was also very good at choosing the best people for his section. Rick, Pieter, Marieke, Annemiek, Mathieu, Geke, Anna, Gael, Erdem, Bia, and Nynke were the people that made my professional life much more fun. I appreciated Bia's attitude towards life. I admired Geke's solidness. Erdem gave me my daily laughs. Gael's world-wide knowledge left me in awe. Rick was by far the best dancer in the group!

I met many people at the coffee corner. Difficult days wouldn't have passed easily if it weren't for the sweet conversations I had with Iemke, Johan, Richard, and Arnold. Marijke D and Annelies could be the big sisters that I never had. CJ's lovely words in

the morning made my day. Bruno’s presence was inspiring. Stella and Marijke M were often the last ones to see me in my long days at the faculty.

StudioLab was another great place where enthusiastic people came together. Pieter Jan Stappers showed me how to think out of the box. Miguel’s taste in food and toys never failed. With Froukje I felt like a girl again. Daniel had the craziest ideas that actually worked. Jasper just talked and it was fun!

I had great time with the DDI girls, Armagan, Sonja, and Linda. Armagan was unique with her vivacious attitude towards life. She was also my connection to Turkey. She had the best Turkish qualities. Boys next to DDI girls were as much fun! Satish and Uri gave the best parties. Satish’s super tasty Indian dishes were something to look forward to. Uri, being larger than life, simply made me happy each time I saw him.

I had a secret affair with the PIM department. I loved working with my direct colleagues all right. But there was something about the PIM people I found appealing. Jan Schoormans always challenged me. Erik Jan always had questions and also answers. With Dirk there was never a dull moment. Being with Katrin and Andre just felt so good. Peter was a-channel-and-two counties away, yet present.

I enjoyed every moment of living in Holland. Piet showed me the quirky side of Holland. Going out with him was a great pleasure. The conversations we had were priceless. Martijn introduced me to the family life in Holland and trusted me his children. Martijntje made me feel special.

Yes, this was a dream come true. In my first day at the faculty, I met a special person who autographed his thesis for me. He wrote “Dear Elif, good luck in Delft”. I had the best luck, Pieterje! You were the inspiration all the way through. Pieter also introduced me to Wim. I sometimes seriously considered that he could be the living Son of God. Having these two beautiful people around me, I couldn’t think better people to be my guardian angels during the defense!

This last act of my life may have been the toughest and the most awarding so far. I did something solely for me. I achieved more that I bargained for. I felt appreciated on the way. All in all, I call it a happy ending.

There is one person that deserves a BIG thank you. While I was busy with myself, he was busy with me too. Alex, Askim, thanks for waiting there and thanks for being here. Amo-te, tanto!

ABOUT THE AUTHOR

⁂

Elif Özcan studied industrial design in Middle East Technical University in Ankara. She worked as a sound designer for a popular radio station (Radyo ODTU) in Ankara, Turkey and as an interaction designer for Novo Design Group in Lisbon, Portugal. In 2003, her academic career started with her PhD study on product sounds in Delft, the Netherlands. At TU Delft, she continues her research on mental processes that underlie sound identification. Her other research interests are developing methods and tools to optimize design processes for product sounds. When she has free time from her research activities, she likes to take photos, manage the sound of a rock band, and play tennis. She speaks Turkish, English, and Portuguese. She is married and dreaming of having her own children.

PUBLICATION LIST

Özcan, E. and van Egmond, R. (2008) Product sound design: An interdisciplinary approach? Proceedings of the 4th International Conference on Design Research. Sheffield: DRS.

Özcan, E., van Egmond, R., and Jacobs, J. (2008) Bases for Categorization and Identification of Product Sounds. In review for JEP: Applied.

Özcan, E. and van Egmond, R. (2008) The effect of visual context on the identification of ambiguous enviornmental sounds. A revision submitted to Acta Psychologica.

Özcan, E. and van Egmond, R. (2007) Memory for product sounds: The effect of sound and label type. Acta Psychologica, doi:10.1016/j.actpsy.2006.11.008

Özcan, E. and van Egmond, R. (2007) How well do we identify product sounds? Proceedings of the 13th International Conference on Auditory Display. Montreal: ICAD.

Özcan, E. and van Egmond, R. (2006) Product sound design and application: An overview. In P.M.A. Desmet, M.A. Karlsson, and J. van Erp (Eds.), Proceedings of the 5th International Conference on Design and Emotion. Gothenburg: Chalmers University of Technology.

Porter, C. S., Desmet, P., Signes, M. J., Sperling, L., Özcan, E., Wouters, C., & Tito, M. (2006) ENGAGE: designing for emotion. In Proceedings of 16th World Congress on Ergonomics. Maastricht: IEA

Özcan, E. and van Egmond, R. (2005) Characterizing descriptions of product sounds. In D Brock (Ed.), Proceedings of the 11th International Conference on Auditory Display. Limerick: ICAD.

Pasman, GJ, Stappers, PJ, Hekkert, PPM, & Keyson, DV (2005). The ID-StudioLab 2000-2005. In H Achten, K Dorst, PJ Stappers, & B de Vries (Eds.), Design research in the Netherlands 2005 (pp. 193-204). Eindhoven: Design systems. (*contributed*)

Özcan, E. and van Egmond, R. (2004) Pictograms for sound design: A Language for the communication of product sounds. In A Kurtgözü (Ed.), Proceedings of the 4th International Conference on Design and Emotion. Ankara: Middle East Technical University.

⁂

aesop's fable - the mice in council

Long ago, the mice had a council to discuss how to outwit their common enemy, the Cat. The mice knew that the Cat was sly and that she always approached them quietly. After long discussions, one young and enthusiastic mouse raised his hand and proposed “all we need is a sign that the Cat is around” and continued “let's put a bell attached by a ribbon round the neck of the Cat. This way, we'll be warned by the bell's jingling sound and have enough time to hide in our holes before the Cat approaches us” This proposal was highly appreciated until an old skeptical mouse asked “but, who is to bell the cat?”

This is a story about impossible solutions. Belling the cat sounds like a crazy idea from the old mouse's conservative perspective. But is it really a crazy idea? The young mouse knows something about sound and sound's function. He realizes sound's effect on the mice and comes up with a proactive solution. What he doesn't know is how to put the bell on the cat. Well, that comes with experience, if he's given a chance to practice his creative ideas.

This thesis is about product sounds. It is about people's responses to the sounds emitted by everyday products. After reading the contents of the thesis, you will know what the young mouse knew. Luckily, you will get some hints about sound design and application. Hopefully, you can convince the old mouse that we can bell the cat!

IDStudioLab

TUDelft