KTH Computer Science
and Communication

Restaurant Soundscapes in Stockholm

A project carried out in collaboration with PerMagnus Lindborg, inspired by the EAT-survey, an investigation of the perceptual quality of servicescapes in Singapore.

EMMA FRID
EMMAFRID@KTH.SE

Project Report at TMH, CSC
Supervisor: Anders Friberg
Examiner: Sten Ternström

2013-07-05

Abstract

Relatively few investigations have yet focused on quality perception of the sonic environment in eateries, i.e. restaurant soundscapes. The aim of this exploratory study is to investigate the correlation between acoustic and perceptual features in such sonic environments. A total of 31 binaural recordings from everyday eateries, divided into the three categories fastfood, café and lunch restaurants, were carried out. On-site annotations of typical restaurant characteristics and perceptual ratings were done on each location, based on the 8-adjectival scale suggested by Axelsson et al. (2010) in the Swedish Soundscape Quality Protocol. Acoustic features were extracted from the recordings and analyzed. Using correlation coefficient analysis, it was found that overall loudness did not significantly correlate with price level of the restaurants. Loudness measures were found to correlate to numerous perceptual features; a louder sonic environment might be perceived as less calm, less pleasant and more chaotic. Mean loudness levels could not be said to significantly differ between the three restaurant categories. The results suggest that the division into three restaurant categories was either substandard, or that the type of restaurant may be of little importance for the actual sonic environment. Furthermore, results indicated that spectral content and soundscape variability may affect perceptual features.

Acknowledgements

I would like to express my very great appreciation to Per-Magnus Lindborg for his guidance and assistance. I would also like to extend my great appreciation to my supervisor Anders Friberg, who has supported and untiringly helped me throughout the entire project process. Special thanks goes to Sten Ternström and Roberto Bresin, who provided me with valuable input and advice in the initial phase of the project.

Table of contents

Chapter 1

Introduction

1.1 Background

The *Handbook for Acoustic Ecology* defines "soundscape" as an "environment of sound with emphasis on the way it is perceived and understood by the individual, or by a society" (Truax, 1999). In soundscape research, focus lies upon the perception and interpretation of an overall acoustic environment. A concept related to the soundscape definition is the notion of "servicescapes", first introduced by Booms & Bitner (1981). The "servicescape" definition emphasizes the role of the physical environment in a situation where a service process may take place, and it has been widely adopted in evaluations of differences in customer experience. Soundscapes such as offices and shopping malls have been rather well researched, and the influence of music in such sonic environments have been thoroughly explored (e.g. Areni & Kim, 1993; Bruner, 1990; Fox 1983; North & Heargraves, 1997, Sterne 1997). However, relatively few investigations have focused on the quality perception of the sonic environment in eating places, i.e. restaurant soundscapes.

Inspired by earlier research within the fields of acoustics and soundscape studies, this study aims to investigate the sonic quality of restaurant environments in Stockholm. The project is mainly inspired by earlier work on restaurant soundscapes carried out by Lindborg (2010-2013).

1.2 Problem Statement

The purpose of this exploratory study is to investigate the sonic quality of restaurant soundscapes in Stockholm by collecting soundscape recordings and data from three different restaurant categories: cafés, lunch- and fastfood restaurants. By extracting acoustic features from soundscape recordings and combining the data with on-site loudness measurements and annotations of architectural properties and perceptual ratings of the sonic environments, the goal is to detect whether there is a sonic difference between restaurant categories. The objective of the study is also to collect recordings and data that can be used in future investigations of the perceptual quality of restaurant soundscapes.

The following questions are to be answered:

- What can be said about the sonic quality and acoustic features characterizing restaurant soundscapes, i.e. servicescapes, in Stockholm? Can we detect any correlations between acoustic and perceptual features in these types of sonic environments?
- By extracting acoustic features from cafés, lunch- and fastfood restaurants, can we detect any correlation between sonic quality and restaurant type?
- How could results and knowledge gained from this study be used within future soundscape research?

1.3 Delimitations

The study can be described as an exploratory survey. The initial project idea was not only to perform acoustic analysis of the soundscape recordings, but to also include a questionnaire that could be distributed among restaurant visitors. The questionnaire was based on the Swedish Soundscape Quality Protocol (2010), thereby enabling perceptual on-site ratings of the sonic environments. Although such a questionnaire was developed and distributed at the first restaurants (see Appendix), the idea of distributing questionnaires on each location was soon left out of the scope of the survey, since this proved to be too time-consuming for the project's limited timeframe. In order to keep the investigation within reasonable limits, perceptual on-site ratings were therefore only annotated by the author. Thorough perceptual rating investigations related to restaurant soundscapes are left as a possible element for future studies within this field, or as a future project collaboration between the author and PerMagnus Lindborg.

Chapter 2

Theory

2.1 Restaurant soundscapes

The soundscape of a restaurant is a dynamic acoustic environment. The sound level in the restaurant does not only depend on the background music level in the environment, but also on the number of restaurant customers and the level of ambient noise created by the diners efforts to maintain a conversation across the tables. Due to the Lombard Effect, i.e. the involuntary tendency of speakers to increase the level of their speech when speaking in noisy environments (Korn, 1954; Webster & Kumpp, 1962), the overall sound pressure in a restaurant will increase even more than the normal additive effect as the number of diners increase (Novak et al., 2010).

The servicescape environment in a restaurant can create a lasting impression for the restaurant visitor. Taste perception can be influenced by auditory cues, such as loudness and background music properties. It has been found that an increase in loudness or beat-per-minutes of background music (i.e. tempo) may result in a significant increase in consumption of food and drink (Zampini & Spence, 2010; Spence & Shankar; 2010, Woods et al., 2011). The servicescape plays an important role in the impression formation of the guests (Bitner, 1992). Moreover, a servicescape may elicit positive or negative reactions that contribute to variability in patronage. (Novak, La Lopa & Novak, 2010)

Apart from the level of background music, noise and loudness levels may also affect a restaurant visitor's experience. As concluded by Kryter (1985), too much sound may result in decreased concentration, increased activity, irritability or tension. A constant sound or silence may nevertheless also be problematic. The key to creating a pleasant sound environment in a restaurant is thus to keep a balance between loud and constant sounds. Due to the fact that the human perception of noise is not absolute in the way that a physical instrument is (such as a sound level meter), the sonic quality of a restaurant cannot be determined by a simple set of loudness measurements alone. The human perception of noise mainly relies on the meaning of the sound in relation to it's sound source and the listener who is exposed to it (Raimbault & Dubois, 2005). Assessments of a sound environment thus depend on the context in which the sounds are perceived (Southwork, 1969). The sounds found in a restaurant should be in tune with the surrounding physical environment (Hellström, 2006).

2.2 Related Work

As concluded by Hence, Baker, Levy & Greval (1992), if management can understand to use the restaurant environment properly, the environment has the potential to become an effective marketing tool. Despite this, relatively few studies have yet focused on investigating restaurant soundscapes (e.g. Lindborg, 2010-2013; Rohrmann, 2003; Novak et al. 2010; Sweetow & Tate, 2000). However, the impact of music on consumer service environments has been thoroughly examined in numerous studies (Milliman 1982, 1986; Yalch & Spangenberg, 1990, 1993; Bruner, 1990; Langrehr, 1991; North & Hargreaves, 1996).

As discussed above, music can be a positive auditory cue that may stimulate consumer emotions and behavior (Lin, 2004). Studies on music and consumer behavior have demonstrated that music may be used as a tool for minimizing negative consequences of waiting in any service operation (Hui, Dube & Chebat, 1997). Moreover, Milliman (1986) found that the tempo of music influences the time customer spends at tables and bars, but that it does not influence service time or purchases of food. Background music may in other words be significant to the results obtained when evaluating a servicescape. Nevertheless, and as stated by Lin (2004), researches should perhaps combine music with other environmental cues when evaluating a servicescape, since humans tend to view an environment holistically when making specific judgments. In order to contribute positively to a restaurant visitor's restaurant experience, it is likely that the background music needs to fit the surrounding restaurant environment.

Novak et al. (2010) used an experimental approach when studying the sonic quality of a fine dining restaurant environment, finding that the presence of appropriate classical music at a "comfortable" level (music plus ambient noise in a range of 62-67 dB(A)) increased the dining pleasure and the overall consumer satisfaction for a clientele in the age span 19-27 years. No music or too loud music was concluded to have negative effects on measures of customer dining pleasure, and women were found to have a greater sensitivity to sound levels than men. The results support the use of music in a restaurant, and suggest that a restaurant with 35-50 patrons (corresponding to an ambient noise level of 53-61 dB(A)) should have a level of music plus ambient noise in the range 62-67 dB(A), with a mean of 64.4 dB(A).

Rohrmann (2003) investigated the sound levels in the eating places and the customers opinions regarding the music situation in these restaurants. The results indicated that restaurant customers have certain preferences and that their satisfaction with a restaurant visit is influenced by the music soundscape in the restaurant. Most customers accepted rather loud sound levels in the restaurants (equivalent loudness levels, Leq, up to 85 dB(A) and peaks above 100 dB(A)). Despite the loudness, such soundscapes were found to be liked or at least tolerated by the restaurant visitors.

In a recent study by Lindborg (2013), 116 various eating-places were investigated. A- and C-weighted equivalent loudness levels (Leq) were estimated and a set of computational acoustic features were extracted from audio recordings. One assumption was that the priciness of the restaurant would influence the sound level. This assumption was, however, not confirmed in the data. The results indicated no significant correlation between the priciness of the restaurants and the ambient noise levels. The SPL levels were persistently higher than the levels that Novak referred to as "highest levels of pleasure and approach behavior". Furthermore, results proved that eating-places in different categories (Chinese, Western, Other Asian) had different sound levels.

Numerous studies investigating soundscapes in Stockholm have recently been carried out (e.g. Axelsson, Nilsson & Berglund, 2010; Nilsson, 2007; Spång et al., 2006). However, most of these have focused on soundscapes in urban areas and parks. Björn Hellström (2006) has made some investigations of the sonic quality of Stockholm restaurants. In the article "Smakade det bra? Va?" Hellström discusses what a good restaurant soundscape actually sounds like. Hellström concludes that a restaurant soundscape should consist of sounds that do not immediately draw our attention, and that restaurant soundscape should be in tune with the surrounding restaurant environment.

2.3 Understanding Soundscapes

2.3.1 Definition

The Canadian composer Murray Schafer first formulated the term soundscape in the late 1960's (1969, 1994). Schafer defined a soundscape as an auditory correspondent of a landscape. The soundscape definition was later extended by Truax, who in the *Handbook for Acoustic Ecology* (1999) defined soundscape as an "environment of sound with emphasis on the way it is perceived and understood by the individual, or by a society". A soundscape can thus be understood as a sonic environment within which we are immersed.

Soundscapes usually contain many sounds that occur simultaneously (Axelsson et al., 2010). Focus in soundscape research lies mainly on the perception and interpretation of an overall acoustic environment. The soundscape can be said to simultaneously be a physical environment and a manner of perceiving this environment. A soundscape not only consists of the physical sounds, but also the material objects that produce these sounds. In general, a soundscape can be said to exist in a particular social and physical context. This close correlation between the interpreter and it's sonic environment places soundscape research close to the field of Acoustic Ecology. Although the soundscape term usually is adopted to refer to actual environments, it is sometimes also used for more abstract or artificial environments or musical compositions.

2.3.2 Acoustic Features

A basic aspect of soundscape investigations is to measure loudness and noise levels. Sound pressure levels are often referred to as the Equivalent Continuous Sound Pressure Level or Leq. Leq is the sound pressure level of an imaginary continuous signal given within a time interval T that would produce the same energy as the real fluctuating sound level with varying amplitude. An Integrating Sound Level Meter with a special weighting characteristic is commonly used for measuring the Leq. A-weighting is the most commonly adopted filter used in surveys evaluating noise in urban soundscapes. By using an A-weighted filter characteristic, a sound level meter can measure the overall sound pressure in a manner that is similar to how the human ear detects sounds. The method is commonly used when measuring broadband sounds, in order to predict how loud they will sound for an average person possessing normal hearing capabilities. (Kryter, 1994)

Apart from A-weighting, there are numerous other filter weighting characteristics available, the C-weighting being one of the most commonly used one. The difference in sensitivity across the frequency range for the different weightings can be seen in Figure 2.1 on the next page. As seen in the figure, the C-weighting is more sensitive to low frequency noise. The C-weighting follows the frequency sensitivity of the human ear at high noise levels, but has a flat frequency response and therefore generally includes more low-frequency components than the A-weighting. When investigating soundscapes, the difference between the estimated A- and C-weighted sound-pressure levels in dB (LeqC-LeqA) can be used as an indicator for the relative proportion of low-frequency noise. (Nilsson, 2007; Nilsson, Andéhn & Lesna, 2008)

Several authors (e.g. Zwicker & Fastl, 1999) have pointed out the limitations of using only A-weighted sound pressure level in sonic evaluations. Zwicker and Fastl developed a complex model of perceived loudness, entitled *Zwicker and Fastl's model for time-varying sounds*. This model is implemented in the Loudness Toolbox for MATLAB, developed by Genesis (2010) and incorporates a model for frequency-dependant loudness perception. Zwicker's `N10` is the loudness exceeded 10 % of the time, a particularly useful measure for overall loudness that have proved to correspond to the perceived loudness. The `N10` percentile, in the unit sone, corresponds to the relative amount of foreground events. The `N90` percentile corresponds to the loudness exceeded 90 % of the time. By calculating `N10` minus `N90`, we get an estimate of the relative prevalence, i.e. the amount of foreground to background events. This may serve as an indicator of soundscape variability, a measure of the relative amount of shorter, louder sound events against the background (Axelsson et al., 2010).

Apart from loudness measures, a number of other acoustic features may be relevant for urban soundscape quality perception. A lot of the knowledge gained from music and emotion research and music information retrieval can also be applied

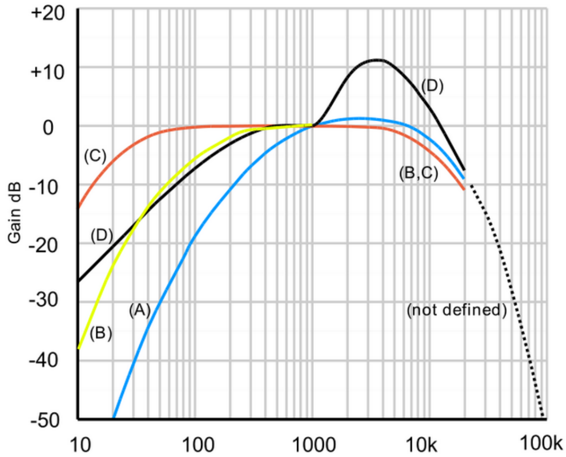

Figure 2.1. A-weighting (blue), B-weighting (yellow), C-weighting (red) and D-weighting (black). (Wikipedia, 2013)

to soundscapes. An important element in music is rhythm, which may express emotional aspects. Juslin and Västfjäll (2008) found that sounds that are sudden, loud, dissonant or feature fast temporal patterns in music may induce arousal or feelings of unpleasantness. Faster rhythms are generally arousing (Juslin & Sloboda, 2010). Rhythm in soundscapes may be related a concept entitled "notice-events", i.e. an instance of an identified and perceived sound (de Counsel, 2009). A rough estimation of the tempo can be detected from periodicities from an onset detection curve. In this particular study, the following parameters related to rhythm are examined: tempo (BPM) and event density. These features are extracted from recordings using the functions `mirtempo` and `mireventdensity`, provided by the MIR Toolbox version 1.3.4 for MATLAB (see Lartillot & Toivainen, 2007).

However, spectral features may also be relevant for soundscape quality perception. The following features, also provided by MIR Toolbox, are investigated in this study: zero-crossing rate (`mirzerocross`), roll-off frequency (`mirrolloff`) and spectral spread (`mirspread` and `mirspectrum`). Zero-crossing rate is an indicator of noisiness. It consists in detecting the number of times a signal changes sign, i.e. crosses the x-axis. The roll-off frequency can be used in order to estimate the amount of high frequencies in a signal. This is done by finding the frequency that the majority (usually 85 or 95%) of the total energy is contained below. The spectral spread is the variance of the Long Time Average Spectrum, LTAS. (Lartillot, 2011) It can be used as a measure of how flat the spectrum is.

2.3.3 Swedish Soundscape Quality Protocol

Nilsson (2007) found that sound source identification within soundscapes may be a better predictor of soundscape quality than actual measured sound levels. In order to measure the on-site perception of urban soundscapes, Nilsson and his colleagues developed the Swedish Soundscape Quality Protocol (SSQP). (Axelsson, Nilsson & Berglund, 2010) SSQP is a framework that can be used in investigations of soundscape quality perception. The framework identifies the underlying dimensions of soundscape perception, thus enabling measurements and improvements of soundscape quality.

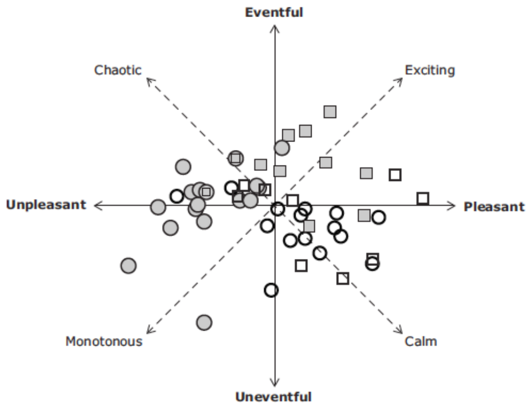

Figure 2.2. The SSQP two-dimensional space. From Lindborg (2013, submitted).

Axelsson et al. (2010) investigated how prerecorded sounds categorized as "human", "natural" and "technological" were perceived. Using excerpts from 50 soundscapes, ratings based on adjectives were collected. Analysis lead to the conclusion that three main dimensions were highly relevant in the perception of soundscapes: Pleasantness, Eventfulness and Familiarity. Urban soundscapes may be represented by their position in a two-dimensional space, defined by the main components Pleasantness and Eventfulness. The two-dimensional space is in turn defined by the following attributes: pleasant, uneventful, exciting, eventful, chaotic, unpleasant, monotonous, uneventful and calm. The SSQP therefore consists of a framework where the quality of a soundscape is rated on these 8 adjectival unidirectional dimensions, represented by 8 vectors added together in a circumplex model (see Figure 2.2).

By rating soundscapes on the mentioned 8 adjective scales, a soundscape can be positioned in the Pleasantness-Eventfulness circumplex with coordinates calculated as follows:

$$X = Pleasantness = \sum R_{Adj} cos(2\pi N_{Adj}/8)$$
$$Y = Eventfulness = \sum R_{Adj} sin(2\pi N_{Adj}/8)$$

Where R_{Adj} is a vector of 8 values [0 - 1] corresponding to the 8 adjectival scales, and N_{Adj} is a whole number [0 - 7] corresponding to the adjective's index in the 8 adjective list.

2.4 Binaural Audio

The soundscape recordings done in this study were all collected using binaural microphones. The term binaural audio refers to the recording and reproduction of sound at the ears. Binaural recordings are carried out by placing a set of miniature microphones close to the ear canals of a human or dummy head. When reproducing the recording in a pair of headphones, the results may be surprisingly realistic. Thus, if using correctly equalized headphones and if there are no head movements during recording or reproduction of the sound, spatial audio may be reproduced by relatively simple means. The limitation of using binaural audio is that all listeners are somewhat different, resulting in that sound signals recorded for a specific subject may not sound entirely correct for another subject. This is due to the fact that all humans will produce individual audio frequency adjustments due to characteristics related to head and ear shape, thereby producing a specific and individual Head Related Transfer Function (HRTF). (Garnder, 2005) The binaural recording-set up used in this investigation can be seen in Figure 2.3.

Figure 2.3. The Sennheiser MK2002 binaural microphones used in the study. (Literaturcafe, 2013)

Chapter 3

Methodology

3.1 Research Design

This investigation has been conducted with an exploratory focus, with the overall goal to explore the correlations between acoustic and perceptual features in restaurant environments in Stockholm. Although the nature of the study made it suitable for quantitative data collection, qualitative data such as detailed on-site annotations were also collected and analyzed. Apart from descriptive data collection and on-site measurements, recordings and photographs (see Appendix) were taken on each site. Evaluation and analysis of data relied mostly on quantitative methods involving statistical and programming tools, such as SPSS and MATLAB.

3.2 Procedure

The selection of sonic environments focused on "partially designed spaces", i.e. environments where some work had been done on the design of the physical space, while at the same time many design aspects have been neglected. The selection of eateries was based on three categories: lunch restaurants, fastfood restaurants and cafés. The selection of these different restaurant categories was founded on the hypothesis that the acoustic features would differ between the category groups. Apart from having different restaurant business concepts, the restaurant categories were assumed to differ somewhat in terms of properties related to the sonic environment, for example in average loudness level.

A total of 31 restaurant soundscape recordings (10 cafés, 10 lunch restaurants and 11 fastfood restaurants) of different indoor eating environments were carried out using a binaural recording system consisting of a set of Sennheiser MKE2002 Binaural Microphones connected to a Zoom H4 recorder. The choice fell on binaural recordings since this would produce spatialization and thus enhance the sensation of presence if the recordings would be used in future listening tests focusing on perceptual ratings. The recording format was set to WAV (44 100 Hz/24 bit) high mic gain and no dynamic processing. The duration of each recording was 5 minutes.

The categories were defined using the following requirements: a "café" would serve mostly salads and sandwiches, not hot meals. Moreover, the occupancy or patronage at this type of restaurant would be somewhat equally distributed during the day. As for the "lunch restaurants", these were defined as restaurants with a very concentrated patronage during lunch hours and a low number of customers after and before these lunch hours. The lunch restaurants also had to serve hot meals and have regular lunch offers. The definition of "fastfood restaurants" was firmly based on their belonging to a set of predefined fastfood chains, such as McDonalds, Burger King and Subway. This category did however also include kebab restaurants and other eateries were the average meal was consumed during a short period of time.

The selected eateries were all situated in the center of Stockholm. Eateries in approximately the same price range were chosen, with a price range spanning from 45 to 130 SEK for an average meal (mean 81 SEK, median 76 SEK). All of the eateries were serving mostly western food. A full list of the eateries can be found in the Appendix. The recordings were evenly distributed during the week, mostly during lunch hours. The median recording hour for respective restaurant category was 14.55 for lunch restaurants, 13.20 for fastfood restaurants and 14.55 for cafés.

Parallel to the audio recordings, on-site SPL measurements using a Center 325 Sound Level Meter were done in order to detect the on-site dB(A) and dB(C). The sound level meter was set to a "fast" time weighting characteristic and ten spot measurements of the peak dB(A) and dB(C) levels were carried out during intervals of 10 seconds. The on-site A- and C-weighted equivalent sound-pressure levels (the LeqA respectively LeqC) were estimated from ten spot measurements (see 4.3.1 Loudness Measures for details). The accuracy of the Center 325 Sound Level Meter was checked both against a Mastech MS6701 Digital Sound Level Meter (SPL levels were checked for different loudness levels of pink noise), and against a 95 dB Sound Calibrator. Results indicated that a correction of + 1.6 dB should be applied in order to provide least RMS-difference between the Center 325 levels and calibrated SPL levels.

As a complement to the audio recordings and SPL measurements, annotations of the descriptive and acoustic features of the restaurants were collected, including observations of size, priciness, occupancy, capacity, background music, characteristics of the visual environment and interior design materials. Photographs of the restaurant were also taken on each recording location. Each soundscape was rated by the author according to the 8 dimensions related to sonic quality perception, as presented in the Swedish Soundscape Quality Protocol. (Axelsson et al. 2010)

After the collection of data and recordings, acoustic features were extracted from each recording using MIR Toolbox (Lartillot, 2011; Lartillot & Toivainen, 2007) and Genesis Loudness Toolbox (Genesis S.A., 2009). For each one of the 31 recordings, the following acoustic features were estimated: zero-crossing rate, roll-off frequency, spectral spread, event density, mirtempo (range constrained to 12-20 BPM, i.e. 0.2-20 Hz) and Zwicker & Fastl's N5, N10 and N90 in sone exceeded 5, 10 respectively 90 % of the time. The difference between sone levels exceeded 10 versus 90 % of the time was also computed, as an indicator for the soundscape variability. Excerpts of 1 minute of the original 5 minute long recordings were used when estimating Zwicker's loudness, since longer excerpts resulted in exceeded available memory the Windows 32-bit system used for extracting the audio features. Apart from Zwicker's loudness measures, the difference between the estimated A- and C- weighted sound pressure levels in dB (LeqC-LeqA) was also calculated and used as an indicator for the relative proportion of low-frequency noise. (Nilsson, 2007; Nilsson, Andéhn & Lesna, 2008). An overview of the acoustic features and their abbrevations can be seen in Table 3.1. The computational features with function names are presented in Table 3.2.

ACOUSTIC FEATURE	*ABBREV.*	*DESCRIPTION*
LeqA	*LeqA*	*Estimated equivalent A-weighted sound pressure level.*
LeqC	*LeqC*	*Estimated equivalent C-weighted sound pressure level. More sensitive to low-frequency noise than LeqA.*
LeqC-LeqA	*CmA*	*Difference between estimated C- and A-weighted SPL. Used as an indicator for relative proportion of low-frequency noise.*
Zero-crossing rate	*zeroX*	*Sign changes across the x-axis, used as an indicator of noisiness.*
Roll-Off	*rolloff*	*The frequency which 85 % oo the total signal energy is contained. Used as an estimation of the amount fo high frequencies in a signal.*
Event density	*eventdens*	*Estimation of the mean frequency of events, i.e. the amount of onsets per second in the temporal data.*
Tempo	*tempo*	*Beats per minute, estimated using onset-detection algorithms.*
Spectral Spread	*spectspread*	*Spectspread: Variance of the LTAS. Used as measure of how flat a spectrum is.*
Zwicker's N5 & N10	*N5, N10*	*Loudness exceeded 5 or 10 % of the time. A measure for overall loudness, corresponding to the relative amount of foreground events.*
Zwicker's N90	*N90*	*Loudness exceeded 90 % of the time. A measure for overall loudness.*
Zwicker's N90-N10	*N10m90*	*Difference in sone levels exceeded 10 versus 90 % of the time. Used as an indicator of soundscape variability, i.e. the amount of foreground to background sonic events.*

Table 3.1. Overview of acoustic features.

FUNCTION	*PARAMETERS*
mirzerocross	*Default*
mirrolloff	*Default*
mireventdensity	*Default*
mirspread(mirspectrum)	*Default*
mirtempo	*Min = 12, Max = 20 BPM*
Loudness_TimeVaryingSound_Zwicker	*x_ratio = 5,10, 90*

Table 3.2. Overview of the functions used for extraction of audio features.

Data analysis was carried out using SPSS (IBM Corp., 2012) and correlation coefficients were computed using a MATLAB script provided by Friberg (see Friberg & Hedblad, 2011 for related work), enabling detection of correlations between acoustic and perceptual features. As a hypothesis test for difference of loudness means, multiple independent-sampled t-tests were used, thereby enabling detection of significant differences in terms of loudness between different restaurant categories. As for the qualitative data analysis, the frequency with which specific key words were used in the annotations of characteristic sounds and materials were calculated, for each separate restaurant category. The purpose of this frequency analysis was to be able to compare characteristic words that were used to describe a specific category with any detected acoustic tendencies related to this type of restaurant.

Chapter 4

Results

4.1 Overview

An overview of all collected data, i.e. the descriptive, perceptual and acoustic features, can be seen in Table 4.2 on the next page. An overview of mean values divided by category can be found in Table 4.1.

	CAFE	FASTFOOD	LUNCH
capacity	69.7	92.8	108.9
occupancy	22.3	46.5	31.4
priciness	86.0	64.3	93.2
overall sonic	2.6	3.5	3.3
overall visual	2.0	3.2	1.8
ambient noise level	2.6	3.2	3.5
characteristic sounds	conversations coffee machine children	background music microwave beeps conversations	conversations background music cutlery
characteristic materials	wooden furniture sofas	glass walls tile floor	wooden floor wooden furniture
LeqA	72.4	73.5	75.3
LeqC	80.2	82.1	79.4
CmA	7.8	8.6	4.2
N5	10.9	12.8	13.2
N10	9.7	11.8	11.9
N90	5.6	7.8	7.0
N10m90	4.1	4.0	4.9
zeroX	956.2	834.2	959.0
rolloff	4137.5	4163.6	4348.1
eventdensity	0.2	0.2	0.3
spectspread	1.2	1.1	1.3
tempo	15.8	15.2	16.0

Table 4.1. Overview of results divided by category. For “overall sonic" and “overall visual", 1 corresponds to “very good’, 2 to “good", 3 to “neutral", 4 to “bad" and 5 to “very bad".

rest	price	capacity	occupacy	LeqA	LeqC	Leq diff	bkg music	music levell	amb. noise	ov. sonic	ov.visual	SSQP1	SSQP2	SSQP3	SSQP4	SSQP5	SSQP6	SSQP7	SSQP8	zeroX	rolloff	eventdens	spectspread	N5	N10	N90	N10m90	tempo
c1	119	100	30	72.40	82.12	9.72	no	can't hear	neutral	neutral	very good	4	2	2	2	4	3	4	4	827.19	3761.41	0.09	1.23	9.90	9.20	6.27	2.93	15.04
c2	80	80	20	62.71	79.01	16.30	no	can't hear	low	good	very good	4	1	3	2	4	1	4	4	914.59	3308.64	0.03	2.09	5.32	4.72	2.32	2.39	15.52
c3	90	200	70	77.35	79.81	2.46	no	can't hear	neutral	neutral	neutral	2	2	3	4	2	2	2	4	1032.80	3679.95	0.29	0.86	15.83	14.37	9.44	4.94	13.91
c4	89	72	13	65.66	77.17	11.51	no	can't hear	low	very good	good	5	1	2	4	5	1	3	2	373.37	3164.92	0.15	0.59	7.48	6.55	3.51	3.04	18.56
c5	69	12	8	75.44	80.95	5.52	no	can't hear	loud	good	bad	4	2	4	2	2	2	5	1	782.90	7156.11	0.02	1.62	10.22	8.79	4.96	3.83	18.67
c6	98	56	25	70.93	87.54	16.61	no	can't hear	neutral	neutral	neutral	3	2	2	3	4	2	3	4	2818.83	4095.98	0.13	1.30	11.35	10.47	7.62	2.86	12.71
c7	89	60	16	73.00	79.08	6.08	yes	OK	neutral	neutral	good	3	4	2	1	1	4	5	2	612.92	4113.29	0.15	0.45	10.91	9.93	6.48	3.45	18.61
c8	89	32	12	72.93	81.90	8.97	no	can't hear	no	good	very good	5	1	2	3	5	2	2	4	493.03	3560.74	0.19	1.70	9.97	8.80	2.93	5.87	15.60
c9	62	35	10	71.47	76.50	5.03	yes	OK	no	good	good	5	2	3	3	4	2	3	3	485.96	4337.82	0.29	0.77	9.93	8.65	3.72	4.93	14.34
c10	75	50	19	81.64	77.83	-3.82	yes	too high	loud	very bad	very good	2	5	4	2	1	4	4	2	1220.03	4196.50	0.26	0.96	17.63	15.77	9.09	6.68	14.90
f1	59	100	20	70.89	82.29	11.39	yes	OK	low	very bad	very good	2	4	5	1	2	3	2	3	777.12	4974.09	0.13	0.62	10.10	8.79	5.29	3.50	17.46
f2	59	48	25	76.27	81.67	5.41	yes	OK	loud	bad	neutral	2	4	3	1	1	4	5	2	791.23	4198.49	0.48	1.30	13.23	12.35	8.13	4.22	16.03
f3	67	88	30	73.44	83.04	9.60	yes	OK	neutral	neutral	neutral	3	3	4	2	1	3	4	4	758.87	3308.64	0.03	0.87	13.90	13.11	8.93	4.18	12.33
f4	45	20	5	66.24	80.72	14.48	yes	OK	low	good	very bad	5	2	2	4	5	2	2	4	537.18	3974.04	0.87	1.51	7.67	7.39	5.80	1.59	19.03
f5	90	270	60	81.13	83.32	2.19	no	can't hear	loud	bad	very bad	2	3	2	4	1	4	4	4	1038.96	3942.73	0.02	0.74	20.31	18.93	12.06	6.87	14.42
f6	67	30	45	72.98	82.49	9.51	yes	OK	loud	very bad	neutral	1	4	2	4	1	5	2	4	793.47	4125.42	0.26	1.07	14.30	13.01	8.42	4.60	13.55
f7	62	70	40	73.82	85.79	11.96	yes	OK	neutral	neutral	good	3	2	2	3	2	2	3	4	793.49	3930.65	0.09	1.13	14.37	12.93	7.98	4.95	15.17
f8	67	150	160	75.74	85.07	9.33	yes	OK	loud	bad	bad	2	2	3	3	2	4	3	4	783.98	3106.88	0.10	0.82	13.53	12.93	8.98	3.95	14.00
f9	69	115	80	77.40	86.89	9.49	yes	OK	loud	neutral	good	2	4	3	4	3	2	3	2	863.26	4151.95	0.43	1.16	12.91	12.03	8.31	3.73	13.14
f10	67	110	40	72.48	76.86	4.38	yes	OK	neutral	neutral	neutral	3	2	2	5	3	4	4	3	1044.12	4650.25	0.16	1.33	11.83	10.68	6.37	4.31	18.64
f11	55	20	6	68.30	75.00	6.70	no	can't hear	low	neutral	bad	4	1	1	4	5	1	2	3	995.08	5436.17	0.01	1.76	8.58	7.81	5.20	2.61	13.43
lr1	70	140	15	63.85	71.68	7.83	yes	too high	neutral	neutral	very good	3	2	2	4	5	2	2	4	787.24	4269.38	0.04	2.44	5.83	5.34	3.04	2.29	19.68
lr2	89	28	10	75.48	79.85	4.37	yes	OK	loud	bad	very good	2	3	2	1	1	5	4	2	1034.41	3789.99	0.28	1.06	13.61	12.53	7.25	5.27	14.04
lr3	75	25	20	79.35	87.72	8.37	yes	OK	extremely loud	bad	good	2	5	4	1	1	4	5	4	888.30	4154.38	1.11	0.82	19.25	17.32	11.97	5.35	15.62
lr4	100	180	20	82.30	82.73	0.43	yes	too low	loud	bad	good	1	5	2	2	1	5	4	4	1077.47	4164.07	0.07	1.15	9.86	8.67	4.71	3.95	16.46
lr5	89	80	7	77.00	84.89	7.89	yes	OK	netural	bad	good	2	2	2	4	1	4	2	2	873.44	4614.79	0.32	1.23	18.24	15.30	6.39	8.91	15.08
lr6	95	150	27	77.22	73.36	-3.86	no	can't hear	low	good	good	5	2	2	4	4	1	2	3	988.51	5637.81	0.43	1.79	12.13	10.54	4.60	5.94	14.61
lr7	98	56	25	69.62	75.59	5.97	no	cant' hear	neutral	neutral	neutral	4	2	2	4	4	2	2	3	918.16	4741.94	0.27	1.63	9.80	9.12	5.12	4.00	12.68
lr8	130	90	40	75.79	77.42	1.63	yes	too low	loud	bad	good	2	4	3	3	2	4	2	4	1087.70	4135.11	0.38	0.97	13.59	13.01	10.03	2.98	18.68
lr9	110	80	75	82.44	87.68	5.24	no	can't hear	loud	neutral	very good	3	2	2	3	2	2	2	5	949.03	4183.79	0.16	0.76	18.09	16.86	10.87	5.99	13.69
lr10	76	260	75	69.52	73.58	4.06	no	can't hear	neutral	good	good	4	3	4	2	2	1	4	4	986.22	3789.99	0.09	1.43	11.37	10.49	6.13	4.36	19.78

Table 4.2. Collected data for each soundscape. Restaurants in the café category are named c1-c10, fastfood restaurants f1-f11 and lunch restaurants lr1-lr10. SSQP1 = pleasant, 2 = chaotic, 3 = exciting, 4 = uneventful, 5 = calm, 6 = annoying, 7 = eventful, 8 = monotonous. For the SSQP perceptual features, 1 corresponds to "Strongly agree", 2 to "Agree", 3 to "Neutral", 4 to "Disagree" and 5 to "Strongly Disagree".

4.2 Descriptive Features

As a complement to the binaural audio recordings and measurements, annotations of the following descriptive features were collected: restaurant size or capacity (i.e. number of chairs), occupancy (i.e. number of visitors), priciness, background music, interior design materials and architecture and characteristic sounds. The complete descriptive data can be found in table 4.2.

OVERALL	N	Range	Min	Max	Mean	Std. Dev.	Variance
Capacity	31	258	12	270	90.55	66.824	4465.389
Occupancy	31	155	5	160	33.81	31.979	1022.628
Av. price	31	85	45	130	80.61	19.366	375.045

Table 4.3. Descriptive features for all restaurants.

As seen in Table 4.1, the priciness was highest in the lunch restaurants, followed by the cafés and lastly the fastfood restaurants. The mean capacity was highest in the lunch restaurants, followed by the fastfood restaurants and lastly the cafés. The mean occupancy was highest for the fastfood restaurants, followed by the lunch restaurants and then the cafés. A number of 18 of the total 31 restaurants played background music, and 13 did not. Only 30 % of the cafés played background music, while the corresponding value for the fastfood restaurants was 82 % and 60 % for the lunch restaurants.

4.2.1 Word frequency

Annotations of the characteristic sounds that could be heard on-site at the restaurants where analyzed using frequency analysis. A transcript including all on-site annotations describing these characteristic sounds and materials can be found in the Appendix.

The words most frequently used in order to describe the characteristic sounds in each restaurant category can be seen in the tables below. The words are organized in descending order, i.e. the words in the first rows are more commonly used in the descriptions than the words in the last rows of the table. According to the frequency analysis, the characteristic sounds are somewhat different in the three categories. Naturally, conversations seems to be common elements in all three categories, and background music was often heard in both fastfood and lunch restaurants, but more rare in the cafés. The sound of a coffee machine was common in the café environments. Beeping sounds from white goods such as microwaves and ovens were usual in the fastfood environments, whereas the sound of cutlery against plates was a common element in lunch restaurants.

SOUND	FREQ.
conversations	5
coffee machine	4
children	3
background music	2
chairs scraping against floor	2
cutlery	2
door opening	2
milk foamer	2
traffic	2
...	

Table 4.4. Characteristic café sounds.

SOUND	FREQ.
background music	6
microwave beeps	5
conversations	4
fridge	3
kebab meat cutter	3
sounds from the till (coins)	3
traffic	3
chairs scraping against floor	2
kitchen sounds	2
rustling paper	2
...	

Table 4.5. Characteristic fastfood sounds.

SOUND	FREQ.
conversations	5
background music	4
cutlery	4
children	4
coffee machine	2
...	

Table 4.6. Characteristic lunch restaurant sounds.

Frequency analysis was also carried out for the collected annotations of the architecture and interior design materials used in the restaurants. As seen in Table 4.7, the most common features that occurred within all three categories were glass walls or facades and tile floors. More wooden materials were generally used in the cafés and lunch restaurants than in the fastfood restaurants. In general, the cafés could be said to be associated to sofas and softer chairs, while only hard materials and hard chairs were used in the fastfood restaurants. In line with the results from the overall visual rating, the lunch restaurants were more commonly associated to the word “decor”.

CAFÉ	FASTFOOD	LUNCH
wooden furniture	glass walls	wooden floor
sofas	tile flooor	wooden furniture
glass walls	linoleum floor	large windows
wooden floor	stone floor	high ceiling
tile floor	no textiles	glass walls
low ceiling	only hard chairs, no sofas	decor

Table 4.7. Characteristic words used to describe interior design & materials. Descending frequency.

4.3 Perceptual features

The selection of perceptual features used for rating of the sonic restaurant environments was founded on the perceptual feature scales presented in the Swedish Soundscape Quality Protocol (Axelsson et al., 2010). The following scales were used for perceptual ratings: overall sonic impression, overall visual impression and degree to which the soundscape could be described as pleasant, exciting, eventful, chaotic, annoying, monotonous, uneventful and calm. Furthermore, the author rated the overall ambient noise level of each restaurant.

4.3.1 Overall Quality

A table of the overall quality divided by category can be seen below. For overall sonic and visual environment, the scale was defined as follows: 1 corresponded to "very good", 2 to "good", 3 to "neutral", 4 to "bad" and lastly 5 corresponded to "very bad". As seen in Table 4.1 and 4.8, the mean overall sonic quality proved to have the highest mean rating in the café category (on average rated from "good" to "neutral"). The second best overall sonic quality rating could be found for the lunch restaurants. The fastfood category received the mean worst mean rating, corresponding to a result between "neutral" and "bad" rating. The overall best sonic quality was found in a quiet café. One café and two fastfood restaurants (McDonalds vs BurgerKing) were judged to have the worst sonic quality.

	CAFE	FASTFOOD	LUNCH
Overall sonic	2.6	3.5	3.3
Overall visual	2.0	3.2	1.8
Ambient noise	2.6	3.2	3.5

Table 4.8. Overall quality of the three categories. For overall sonic and visual environment, 1 corresponds to "very good" and 5 to "very bad". For ambient noise level 1 corresponds to "no noise" and 5 to "extremely loud noise".

As for the visual environment, the mean rating was highest for lunch restaurants (on average considered as "good"), followed by cafés and lastly by fastfood restaurants. The mean ambient noise level in the restaurants was 3.1 on a scale from 1 to 5, where 1 corresponding to no noise and 5 to extremely loud noise. The mean ambient noise level was rated highest in the lunch restaurant category (3.5), followed by the fastfood restaurant category (3.2) and lastly the cafés (2.6).

4.3.2 Qualia

The 8 different adjectives of SSQP creates a circumplex model for soundscape quality perception, introduced by Axelsson, Nilsson and Berglund (2010). A plot of the quality ratings, i.e. the Qualia ratings, for the 31 different soundscapes investigated in this study can be seen in Figure 4.1, where different restaurant categories have been plotted using dots of different color. As seen in the figure, no clear pattern can be distinguished for respective restaurant category. Soundscapes from similar categories do not appear in any specific noticeable clusters. The quality perception of restaurants seem to vary to a certain extent within the three restaurant categories, according to the author's Qualia ratings.

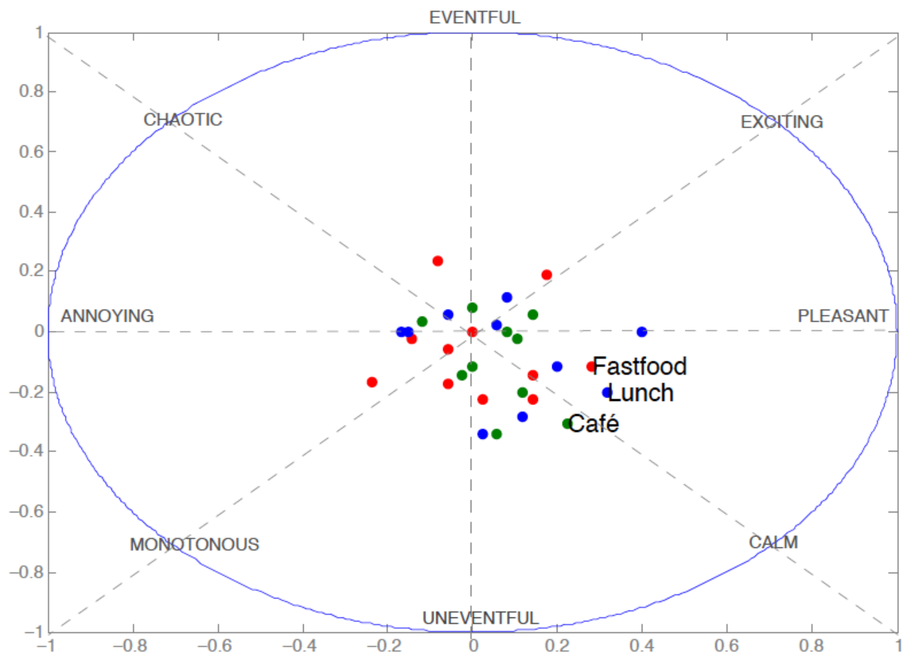

Figure 4.1. Circumplex projection of soundscape quality ratings according to the SSQP model. The soundscapes have been coded in colors depending on restaurant category: red dots correspond to fastfood restaurants, green to cafés and blue to lunch restaurants.

4.4 Acoustic Features

4.4.1 Loudness Measures

A- and C-weighted equivalent sound pressure levels (L_{Aeq} respectively L_{Ceq}) were estimated from 10 peak dB spot measurements done during the binaural recordings (10 spot measurements during 2.5 minutes for each weighting characteristic). The following formula was used for calculating the estimated equivalent sound pressure level:

$$10 \log_{10} \frac{\sum_{n=1}^{n} 10^{\frac{Lp_n}{10}}}{n}$$

Where Lp_n is the relative sound pressure of sample n. In our case, $n = 10$. The difference between A - and C - weighted equivalent sound pressure level was also computed for each restaurant. In order to estimate Zwicker's loudness Nx, Genesis Loudness Toolbox (Genesis S.A., 2009) with default settings was used. A summation of all of the loudness measures, presented with a +1.6 dB correction, can be seen in the Table 4.9. As seen in the table, the LeqA ranged from 62.7 to 82.4 dB.

	Mean	Min	Max	Std. Dev.	Variance
LeqA	73.7	62.7	82.4	5.2	26.6
LeqC	80.6	71.7	87.7	4.4	19.6
CmA	6.9	-3.9	16.6	4.9	24.2
N5	12.3	5.3	20.3	3.8	14.3
N10	11.2	4.7	18.9	3.5	12.1
N90	6.8	2.3	12.1	2.6	6.8
N10m90	4.3	1.6	8.9	1.6	2.4

Table 4.9. Overview of the loudness features: equivalent sound pressure level (dB SPL) and Zwicker's loudness Nx (sones).

The loudness features organized by category, applied with 1.6 dB correction, can be seen in the Table 4.13. The mean LeqA levels were 72.4 dB for cafés, 73.5 dB for fastfood restaurants and 75.3 dB for lunch restaurants. In order to detect whether or not the loudness data actually differed between restaurant categories, multiple independent-samples t-tests were carried out. A t-test is generally used for deciding whether a difference between two categories of data is significant, i.e. if there is an actual difference in mean between two variables. In this case, the t-test asks wheter the mean difference is representative for the difference between restaurant categories in general, or if it is likely to have been produced by random chance in the sample selection. The difference in mean can be said to be representative of the real difference between two categories, if the t-test proves to be significant. By performing multiple t-tests for the loudness features LeqA and LeqC, it was found that the three restaurant categories can not be said to be different in terms of equivalent sound pressure level.

		Mean	St. Dev.	Variance
Cafe	LeqA	72.4	5.4	29.1
	LeqC	80.2	3.2	10.3
	CmA	7.8	6.2	38.7
	N5	10.9	3.6	12.9
	N10	9.7	3.3	10.8
	N90	5.6	2.6	6.5
	N10m90	4.1	1.4	2.1
Fastfood	LeqA	73.5	4.2	17.4
	LeqC	82.1	3.6	12.8
	CmA	8.6	3.6	12.9
	N5	12.8	3.4	11.5
	N10	11.8	3.2	10.3
	N90	7.8	2.0	4.1
	N10m90	4	1.3	1.8
Lunch R.	LeqA	75.3	6.0	35.6
	LeqC	79.4	6.0	36.4
	CmA	4.2	3.9	14.9
	N5	13.2	4.3	18.6
	N10	11.9	3.8	14.7
	N90	7.0	3.0	8.9
	N10m90	4.9	1.9	3.5

Table 4.10. Overview of the loudness features for respective restaurant category.

4.4.2 Other computational audio features

Computational audio features were extracted from the binaural audio files using MIR Toolbok (Lartillot, 2011). An overview of the computational audio features can be seen in the table below:

	Unit	Min	Max	Mean	Std. Deviation	Variance
zeroX	Hz	373.37	2818.83	913.83	403.88	163118.16
rolloff	Hz	3106.88	7156.11	4214.71	793.84	630189.56
eventdens	Hz	0.01	1.11	0.24	0.24	0.06
spectspread	Hz	0.45	2.44	1.20	0.46	0.21
tempo	BPM	12.33	19.78	15.66	2.28	5.18

Table 4.11. Overview of the computational features extracted using MIR Toolbox.

4.5 Cross-correlation comparison

In order to investigate if there are any correlations between perceptual and acoustic features in restaurant soundscapes, pairwise cross-correlation coefficients were analyzed. Using the provided correlation coefficient function `[R,P]=corrcoef(...)` in MATLAB, a correlation coefficient matrix R and p-value matrix P can be computed. The p-values are testing the hypothesis of no correlation between variables.

Each p-value in the P-matrix will correspond to the probability of getting a correlation of the same size as the observed value by randomized chance, if the true correlation is zero. If the p-value is small, usually $p < 0.05$, (alternatively $p < 0.01$ or $p < 0.001$) the correlation $R(i, j)$ can be said to be significant. (Matlab, 2013)

By analyzing the values of the p-value matrix P and the correlation coefficient matrix R, conclusions about the correlation between computational acoustic features and perceptual concepts can be drawn. The first step of the analysis of correlation coefficients consists of sorting out which features that are not significantly correlated in the matrix P. The next step is to consider the size of the remaining significant correlations, i.e. to investigate the correlation coefficients in matrix R. According to Williams (1968) a correlation coefficient between 0.7-0.9 should be considered a marked relationship, whereas a correlation coefficient between 0.4-0.7 should be considered a substantial relationship. The sorting into marked and substantial relationships is useful when interpreting the relationship between variables, since high correlations reveal dependencies (Hedblad, 2011).

4.5.1 Interdependence

Perceptual and descriptive features

The interdependence of different perceptual rating scales and descriptive features was investigated using cross correlations. Results are shown in Table 4.12. The pairwise correlation coefficients, r, are marked with one to three stars depending on their level of significance ($* = p < 0.05, ** = p < 0.01, *** = p < 0.001$). The rating scales for the perceptual features follow the following pattern: ambient noise level 1-5 (1: no noise ... 5: extremely loud noise) overall sonic and visual 1-5 (1: very good ... 5: very bad) and SSQP adjectival scale 1-5 (1: strongly disagree ... 5: strongly agree).

Less than half of the correlations in Table 4.12 can be said to be significant (p<0.05). As seen in the table, numerous of the r coefficients have both high values plus a level of significance below 0.001. Ten correlations have $r > 0.7$ and can therefore be said to be marked relationships according to Williams (1968). The "ambient noise level" negatively correlates to "pleasant" and "calm". The "overall sonic impression" correlates negatively to "pleasant" and positively to "chaotic" and "annoying", since these scales are in reversed order. The "pleasant" feature is negatively correlated both to "chaotic" and "annoying" but positively correlated to "calm". Furthermore, "calm" is negatively correlated to both "chaotic" and "annoying". There is a substantial correlation between "uneventful" and "eventful" ($r = -0.63$), which might indicate that the author has interpreted these words somewhat as opposites, which is not entirely correct, since the words actually have a certain degree of semantical difference.

Surprisingly, there does not seem to be any correlation between average price and ambient noise rating, or overall sonic impression. The overall visual environment is only correlated to the average price on a significance level of $p < 0.05$. Furthermore, no correlation between capacity and ambient noise level or overall sonic impression can be detected.

	average price	capacity	occupacy	ambient noise	SSQP ov. sonic	SSQP ov. visual	SSQP pleasant	SSQP chaotic	SSQP exciting	SSQP uneventful	SSQP calm	SSQP annoying	SSQP eventful
capacity	0.23												
occupacy	0.05	0.54**											
ambient noise	0.19	0.13	0.37*										
SSQP ov. sonic	0.01	0.02	0.16	0.59***									
SSQP ov. visual	-0.36*	0.09	0.20	0.11	-0.05								
SSQP pleasant	-0.07	-0.21	-0.32	-0.75***	-0.87***	0.02							
SSQP chaotic	0.02	0.08	0.04	0.64***	0.70***	-0.13	-0.71***						
SSQP exciting	-0.24	0.08	0.14	0.19	0.21	-0.15	-0.20	0.45*					
SSQP uneventful	-0.02	0.19	0.16	-0.25	-0.26	0.37*	0.21	-0.49**	-0.54**				
SSQP calm	-0.05	-0.18	-0.24	-0.71***	-0.68***	-0.01	0.78***	-0.72***	-0.40*	0.47**			
SSQP annoying	0.11	-0.03	0.07	0.62***	0.79***	0.02	-0.77***	0.67***	0.05	-0.30	-0.73***		
SSQP eventful	-0.08	0.00	-0.07	0.42*	0.04	0.03	-0.15	0.39*	0.33	-0.63***	-0.46**	0.33	
SSQP monotonous	0.22	0.34	0.35	-0.03	0.02	-0.01	-0.04	-0.11	-0.13	0.15	0.12	-0.09	-0.33

Table 4.12. Inter-correlations between perceptual and descriptive ratings. N=31, p-values: * < 0.05; ** < 0.01, ***<0.001. Marked relationships are highlighted with a cyan color.

	LeqA	LeqC	CmA	zeroX	rolloff	eventdens	spectspread	N5	N10	N90	N10m90
LeqC	0.48**										
CmA	-0.62***	0.39*									
zeroX	0.15	0.23	0.05								
rolloff	0.12	-0.21	-0.31	0.06							
eventdens	0.17	0.19	-0.01	-0.11	-0.01						
spectspread	-0.48**	-0.45*	0.10	0.08	0.29	-0.10					
N5	0.82***	0.52**	-0.39*	0.17	-0.06	0.26	-0.54**				
N10	0.81***	0.52**	-0.37*	0.18	-0.11	0.26	-0.55**	0.99***			
N90	0.68***	0.55*	-0.21	0.26	-0.18	0.30	-0.57***	0.87***	0.91***		
N10m90	0.67***	0.25	-0.48**	-0.02	0.06	0.08	-0.27	0.77***	0.72***	0.36*	
tempo	-0.33	-0.42*	-0.03	-0.31	0.14	0.05	0.14	-0.39*	-0.40*	-0.33	-0.33

Table 4.13. Inter-correlations between computed acoustic features. N=31, p-values: * < 0.05; ** < 0.01, ***<0.001. Marked relationships are highlighted with a cyan color.

Acoustic features

The pairwise correlations between the extracted audio features can be seen in Table 4.10. Less than half of the correlations were significant. Five marked relationships can be detected, all of them being relationships between loudness features. As expected, LeqA, LeqC and the Zwicker's Nx features correlate rather well, since they are all loudness measures. Naturally, multiple high correlation coefficients can be seen between different Zwicker's measures. Overall, the correlation coefficients related to loudness indicates that many of the features work in the way they are intended, since similar measures have high correlations. As opposed to what was expected, the event density does not significantly correlate with the tempo feature.

This might indicate a problem, since both methods are based on on-set detection and should be related to some extent. Nevertheless, this might be explained by the fact that the event density function is not fully reliable when it comes to predicting musical tempo. Thus it is maybe not too surprising that this feature is even less exact when it comes to predicting tempo for soundscape recordings. Moreover, the spectral features such as rolloff and spectspread do not correlate significantly either. One could have assumed that these features would correlate, since they are both related to the relative amount of high-frequency content in the signal.

4.5.2 Comparison

In this section, the perceptual and descriptive features are compared to the extracted audio features. The pairwise correlations between the ratings are presented in Table 4.14. Out of the 14 perceptual and descriptive features, only 2 had marked relationships ("SSQP calm" and "ambient noise level") with acoustic features. In total, only 20 % of the correlations were significant ($p < 0.05$ or less) and only three marked relationships could be found: N5 and "calm" ($-r = 0.72$), N10 and "calm" ($r = -0.71$) and N90 and ambient noise level ($r = 0.74$). These results would indicate that an increase in Zwicker's loudness (i.e. in sonic foreground events) would result in a soundscape perceived as less calm. An increase in overall N90 Zwicker's loudness would produce a higher perceived ambient noise level.

	LeqA	LeqC	CmA	zeroX	rolloff	eventdens	spectspread	N5	N10	N90	N10m90	mirtempo
average price	0.31	0.04	-0.28	0.29	-0.14	-0.14	-0.13	0.18	0.19	0.15	0.16	-0.09
capacity	0.21	-0.14	-0.34	0.10	-0.23	-0.31	-0.06	0.19	0.20	0.18	0.16	0.11
occupacy	0.30	0.29	-0.05	0.06	-0.34	-0.14	-0.27	0.34	0.39*	0.48**	0.08	-0.22
ambient noise	0.62***	0.42*	-0.27	0.25	0.02	0.20	-0.27	0.61***	0.64***	0.74***	0.20	-0.10
SSQP ov. sonic	0.51**	0.33	-0.23	0.22	-0.02	0.09	-0.34	0.55**	0.55**	0.55**	0.31	-0.25
SSQP ov. visual	0.01	0.08	0.06	0.07	0.20	0.09	-0.03	0.10	0.13	0.26	-0.14	-0.07
SSQP pleasant	-0.57***	-0.44*	0.19	-0.26	0.13	-0.04	0.37*	-0.57***	-0.58***	-0.62***	-0.26	0.20
SSQP chaotic	0.55**	0.23	-0.37*	0.09	-0.01	0.35	-0.43*	0.43*	0.44*	0.50**	0.15	0.05
SSQP exciting	0.14	0.10	-0.05	-0.07	0.07	0.14	-0.26	0.17	0.16	0.19	0.03	0.15
SSQP uneventful	-0.17	-0.21	-0.01	0.01	0.06	-0.06	0.27	-0.05	-0.05	-0.09	0.05	-0.10
SSQP calm	-0.69***	-0.41*	0.36*	-0.06	0.03	-0.06	0.56**	-0.72***	-0.71***	-0.66***	-0.50**	0.12
SSQP annoying	0.54**	0.33	-0.27	0.04	-0.14	0.15	-0.44*	0.49**	0.49**	0.47**	0.31	-0.05
SSQP eventful	0.21	0.15	-0.09	0.02	-0.01	0.00	-0.20	0.12	0.13	0.17	-0.00	0.18
SSQP monotonous	-0.04	0.19	0.21	0.17	-0.47**	-0.02	0.07	0.07	0.12	0.23	-0.13	-0.15

Table 4.14. Cross-correlations between computed acoustic features and perceptual ratings. N=31, p-values: * $<$ 0.05; ** $<$ 0.01, ***$<$0.001. Marked relationships are highlighted with a cyan color.

A number of substantial correlations can be found. Going from left to right, we can see that the LeqA level is significantly ($p < 0.001$) correlated to the ambient noise level ($r = 0.62$) and the perceptual features "pleasant" ($r = -0.57$) and "calm" ($r = -0.69$). This implies that an increase in LeqA would result in an increase in perceived ambient noise level, and a lower rating on the "pleasant" and "calm" scale. On the significance level of $p < 0.01$, we can detect the following correlations to LeqA: “overall sonic” ($r = 0.51$), ’chaotic’ ($r = 0.55$) and “annoying” ($r = 0.54$). This indicates that the overall sonic rating and the perceptual features “chaotic” and “annoying” would obtain higher values, i.e. worse ratings, for increased LeqA levels. As for the acoustic features LeqC and CmA, the relative proportion of low frequency noise, only correlations with perceptual features on a significance level of $p < 0.01$ can be detected.

No significant correlations at all can be found between zero-crossing rate, i.e. noisiness, and the perceptual features. As for roll-off frequency, only one significant correlation can be found; a relationship to the parameter “monotonous” ($r = -0.47$). This substantial relationship indicates that a soundscape with a higher roll-off frequency would be perceived as less monotonous. For the feature event-density, no significant correlations can be found. The spectral spread, i.e. the variance of the LTAS spectrum, substantially correlates to the perceptual feature “calm”, indicating that an increase in spectral spread would also increase the “calm” rating.

As seen in Table 4.14, the Zwicker’s Nx parameters are correlated to a number of perceptual features. There is a marked relationship between the parameters N5 and N10 and the perceptual feature “calm” ($r = -0.72$ and $r = -0.71$), indicating that an increase in loud foreground sounds would reduce the “calm” rating. Furthermore, substantial relationships exist between N5, N10, N90 and the parameters “ambient noise level” ($r = 0.61 - 0.74$), “overall sonic impression” ($r = 0.55$), “pleasant” ($r = -0.62 - 0.57$) and “annoying” ($r = 0.47 - 0.49$). Thus, an increase in foreground events may result in increased ambient noise level, a lower “pleasant” rating and an overall worse sonic impression, since this rating scale is reversed to the adjectival scales. As for the N90 feature, a marked relationship (r =0.74) can also be found to the ambient noise level and substantial relationships to the number of guests, as well as the feature “chaotic”. Only one significant correlation can be found in the N10m90 column; a substantial relationship to the “calm” feature ($r = -0.50$) indicating that the soundscape would be perceived as less “calm” if the soundscape variability would increase. Lastly, no correlations at all can be found between the tempo feature and perceptual ratings.

Quite surprisingly, no correlations between average price or capacity and loudness features can be found in the cross-correlation table. As for the number of guests, a substantial correlation can only be found for the N90 feature. Despite the fact that a restaurant visit is usually characterized by multiple sensory stimulations, the overall visual impression does not seem to be affected by any acoustic features, or vice versa.

Chapter 5

Discussion

5.1 Discussion and conclusion

As suggested by Novak et al. (2010), in a restaurant with 35-50 patrons, the level of music plus ambient noise should be in the range of 62-67 dB(A) with a mean of 64.4 dB(A). The mean equivalent sound pressure levels for all restaurant categories investigated in this survey exceed these dB(A) levels. The range in equivalent sound pressure level for all restaurants was found to be 63-82 dB(A), with an overall mean level of 73.7 dB. This suggests that a lot can be done in order to improve the sonic quality of Stockholm restaurants, considering loudness levels. When analyzing photographs, annotations of the materials used and architectural design of the different restaurants, one can conclude that much of the noise in restaurant environments may come from echo and reverberation due to the frequent use of large glass walls, tile floors and lack of carpets, soft chairs, textiles and other natural silencers. However, as stated by Rohrmann (2003), restaurant customers may accept sound levels and Leq up to 85 dB(A). However, according to the regulations of Arbetsmiljöverket (2005) an equivalent noise level of 80 dB(A) should not be exceeded on a workplace if the workday is eight hours. If a level of 85 dB(A) is measured on an industrial workplace, hearing protection should be offered. It is thus rather remarkable that restaurant visitors accept such high sound levels.

To conclude, although rather high sound levels were found in the majority of the restaurants investigated in this survey, this might not be perceived as a problem for the average restaurant visitor, and does not necessarily have to affect the overall dining experiences. A loud sound level may not be physically annoying, if it happens to have a suitable frequency distribution.

As found in earlier studies by Lindborg (2010-2013), the results of this investigation indicate no significant correlation between priciness of restaurants and the ambient noise levels or overall loudness. Surprisingly, the average price level of a restaurant does not seem to correlate to any acoustical or perceptual feature. The most prominent correlations were found for loudness measures, such as LeqA and Zwicker's N10. The loudness measures were found to be related to a number of perceptual features, such as overall ambient noise level, overall sonic impression and perceptual rating of “calm”, “pleasant” and “chaotic”. According to the results, increased loudness would lead to the perception of a less calm, less pleasant and more chaotic sonic environment. Moreover, the following substantial relationships were found; a higher roll-off frequency seemed to lead to sonic environment perceived as less “monotonous”, an increase in spectral spread would improve the rating of the perceptual feature “calm”, and an increase in soundscape variability (N10m90) would reduce the “calm” rating.

The division into different restaurant categories done in this study proved to have little effect on the overall loudness level in the restaurant. Although the mean loudness levels, as well as other acoustic features, varied between the different categories, it is not possible to claim that a significant difference can be found in loudness for the three categories. The conclusion that the categories are not easily separated applies not only for the loudness levels, but also for perceptual features, as seen when plotting the quality ratings of each restaurant type in the SSQP circumplex. The results suggest that the division into three restaurant categories was either substandard, or that the type of restaurant may be of little importance for the actual sonic environment. It is possible that other properties than restaurant type, i.e. restaurant business concept, are more important for the overall acoustic and perceptual quality of a restaurant. Nevertheless, some differences concerning typical sounds heard in and typical interior materials used in the respective restaurant category were found. For example, wooden materials seem to be more commonly used in cafés and lunch restaurants, and the fastfood restaurants often lack of decor or softer materials. The price range, capacity and occupancy also differed between the restaurant types, however, this seemed to have little effect on the acoustic or perceptual features.

5.2 Improvements and future work

Due to the exploratory nature of this study, the obtained results should be considered only preliminary. Morever, the perceptual ratings used in the cross-correlation analysis were founded on single ratings done by the author, not on ratings done by multiple subjects. For proper investigations involving evaluations of perceptual features in soundscapes, perceptual ratings would preferably have to been done by at least 5 to 10 subjects on each recording location. The relative strength of correlations found will have to be examined in future investigations and research.

Although many acoustic features that are usually examined in research related to music information retrieval can also be utilized when investigating soundscapes, the tempo functions used in this survey, i.e. those provided by MIR Toolbox, are probably better suited for music tempo detection than for tempo detection in soundscapes. The selection of acoustic features used for analysis of the soundscapes done in this investigation could perhaps have been improved, in order better suit analysis of sonic restaurant environments.

A suggested continuation of this study could be to use the collected binaural recordings in a listening experiment focusing on the perceptual features presented in the Swedish Soundscape Quality Protocol (Axelsson et al., 2010). Since similar studies and data collections have already been carried out in Singapore (e.g. Lindborg, 2010-2013; Lindborg, 2013, submitted), the data collected in this Stockholm investigation could maybe be used in a cross-cultural study. Such a study would involve perceptual ratings of sonic restaurant environments originating from different geographical contexts, and also subjects originating from different countries.

Another possible extension of this study could be to further investigate whether audio features can actually predict perceptual features of restaurant soundscapes, i.e. if it would be possible to predict how a sonic environment of this type would be perceived by analyzing acoustic features. If correlations between acoustic and perceptual features could be predicted to such a extent, one would be able to conclude approximately how a restaurant environment would be perceived by potential customers. Thus, knowledge gained from this specific research field may not only be of scientific interest, but also have possible commercial applications.

Bibliography

Arbetsmiljöverket. (2005) Buller - Arbetsmiljöverkets föreskrifter om buller samt allmänna råd om tillämpningen av föreskrifterna. http://tinyurl.com/pt33qc2 Accessed on 2013-11-23.

Areni, C. S., & Kim, D. (1993). The Influences of Background Music on Shopping Behavior. Advances in Consumer Research. 20(1). pp 336-340.

Axelsson, Ö., Nilsson, M.E. & Berglund, B. (2010) A Principal Coponents Model of Soundscape Perception. Journal of the Acoustical Society of America. 128(5). pp 2836–2846.

Bitner, M.J. (1992) Servicescapes: The Impact of Physical Surroundings in Customers and Employees. Journal of Marketing. 56. pp 57-71.

Booms, B.H. & Bitner, M.J. (1981). Marketing Strategies and Organisation Structures for Service Firms. In Donnelly, J; George, WR. Marketing of Services. Chicago, IL: American Marketing Association.

Bruner, G. (1990) Music, Mood and Marketing. Journal of Marketing. 54(4). pp 94-104.

de Coensel, B. (2009) A Model for the Perception of Environmental Sound Based on Notice-Events. Journal of the Acoustical Society of America. 126(2). pp. 656-665.

Friberg, A. & Hedblad, A. (2011) A Comparison of Perceptual Ratings and Computed Audio Features. In 8th Sound and Music Computing Conference, Padova, Italy.

Gardner, W.G. (2005) Spatial Audio Reproduction: Toward Individualized Binaural Sound. In Frontiers of Engineering. Alington, Massachusetts. pp 114-115.

Genesis S.A. (2009) Loudness Toolbox. For Matlab.
http://www.genesis-acoustics.com/ Accessed on 2013-06-20.

Hedblad, A. (2011) Evaluation of Musical Feature Extractio Tools using Perceptual Ratings. Master of Science Thesis, KTH, Stockholm.

Hellström, B. (2006) Smakade det bra? Va? Martins Magasin. 3/4-28/4 2006. pp 42-45. http://www.acousticdesign.se/public/article.asp?id=23 Accessed on 2013-06-26.

Hui, M.K., Dube, L. & Chebat, J. (1997) The Impact of Music on Consumers' Reactions to Waiting for Services. Journal of Retailing. 73(1). pp 87-104.

IBM Corp. (2012) IBM SPSS Statistics for Mac, Version 20.0. Armonk, NY: IBM Corp.

Juslin, P. & Sloboda, J. (2010) Handbook of Music and Emotion. Oxford Press.

Juslin, P. & Västfjäll, D. (2008) Emotional Responses to Music: The Need To Consider Underlying Mechanisms. Behavioural and Brain Sciences. 31(5). pp 559-621.

Korn, T. S. (1954) Effect of Psychological Feedback on Conversational Noise Reduction in Rooms. Journal of the Acoustic Society of America. 26(5). pp 793–794.

Kryter, K.D. (1985) The Effect of Noise on Man. Academic Press. Orlando, Fl. 2nd Ed.

Kryter, K.D. (1994) The Handbook of Hearing and the Effects of Noise. Physiology, Psychology and Public Health. San Diego, CA: Academic Press.

Lartillot, O. & Toivainen, P. (2007) A Matlab Toolbox for Musical Feature Extraction. From Audio. In: International Conference on Digital Audio Effects. 2007.

Lin, I.Y. (2004) Evaluating a Servicescape: The Effect of Cognition and Emotion. International Journal of Hospitality Management. 23(2). pp 163-178. `http://www.sciencedirect.com/science/article/pii/S0278431903000951` Accessed on 2013-06-16.

Lindborg, PM. (2010-13). Soundscape Emotion Response Studies `http://www.permagnus.net/pm/research/SSER/index.html` Accessed on 2013-04-15.

Lindborg, PM. (2013, submitted) Perception of Soundscapes Correlates with Acoustic Features and is Moderated by Personality Traits.

Matlab (2013) Correlation Coefficients. Matlab rR2013a Documentation. `http://www.mathworks.se/help/matlab/ref/corrcoef.html` Accessed 2013-07-03.

Milliman, R.E. (1982) Using Background Music to Affect the Behaviour of Supermarket Shoppers. Journal of Marketing. 46(3). pp 86-91.

Milliman, R.E. (1986) The Influence of Background Music on the Behaviour of Restaurant Patrons. Journal of Consumer Research. 13(2). pp 286-289.

Nilsson, M.E. (2007a) Soundscape Quality in Urban Open Spaces. In Proceedings of INTER-NOISE 2007, Istanbul, Turkey.

Nilsson, M.E. (2007b) A-weighted Sound Pressure Level as an Indicator of Perceived Loudness and Annoyance of Road-traffic Sound. Journal of Sound and Vibration. 302(1-2). 197-207.

Nilsson, M.E., Andéhn, M. & Lesna, P. (2008) Evaluating Roadside Noise Barriers Using an Annoyance-Reduction Criterion. Journal of the Acoustic Society of America. 124(6). 3561-3567.

North, A.C. & Hargreaves, D.J. (1996) The Effects of Music on Response to a Dining Area. Journal of Environmental Psychology. 16(1). pp 55-64.

North, A.C. & Hargreaves, D.J. (1997) Music and Consumer Behaviour. In A.C. North & D.J. Hargreaves (Eds.), The Social Psychology of Music. New York: Oxford Univeristy Press.

Novak, C.C., Lopa, J. & Novak, R.E. (2010) Effects of Sound Pressure Levels and Sensitivity to Noise on Mood and Behavioural Intent in a Controlled Fine Dining Restaurant Environment. Journal of Culinary Science & Technology. 8(4). pp 191-218.

Raimbalt, M. & Dubois, D. (2005) Urban Soundscapes: Experiences and Knowledge. Cities. 22(5). pp 339-350.

Schafer, R.M. (1994) The Soundscape: Our Sonic Environment and the Tuning of the World Destiny Books. Rochester, VT.

Schafer, R.M. (1969) The New Soundscape Associated Music. New York, NY. pp 1-65.

Southwork, M. (1969) The Sonic Environment of Cities. Environment and Behavior. 1(1). pp 49–70.

Spence, C. & Shankar, M.U. (2010) The Influence of Auditory Cues on the Perception of, and Responses to, Food and Drink. Journal of Sensory Studies. 25(3). pp 406–430.

Sweetow, R. & Tate, L. (2000) I'll have a side order of earplugs, please. Audiology Today. 12(3). p 40.

Thompson, E. (2004) Soundscape of Modernity. Architectural Acoustics and the Culture of Listening in America, 1900-1933. Chap. 1. MIT Pess.

Yalch, R. & Spangenberg, E. (1990) Effects of Store Music on Shopper Behaviour. Journal of Services Marketing. 4(1). pp 31-39.

Spång, K. Öhrström, E., Gidlöf-Gunnarsson, A., Berglund, B., Nilsson, M., Forssén, J. Hornikx, M., Kropp, W. Kihlman, T. (2006) Soundscape Support to Health - Summaries of studies and results. `http://www.ljudlandskap.acoustics.nu/downloads/publikationer/ljudlandskap/slutrapporter/Summary+of+studies+and+results.pdf` Accessed on 2013-06-25.

Sterne, J. (1997) Sounds Like the Mall of America: Programmed Music and the Architectonics of Commercial Space. Ethnomusicology. 41(1). pp 22-50.

Thompson, E. (2002) The Soundscape of Modernity. Cambridge: MIT Press.

Truax, B. (1999) Handbook for Acoustic Ecology, 2nd ed. `http://www.sfu.ca/sonic-studio/handbook/`. Accessed on 2013-06-16.

Truax, B. (2011) Acoustic Tradition and the Communicational Approach. Acoustic Communication, 2nd ed. Ablec, Westport, CT. pp 3-14.

Woods, AT., Poliakoff, E., Lloyd, DM., Kuenzel, J., Hodson, R., Gonda, H., Batchelor, J., Dijksterhuis, G.B., & Thomas, A. (2011) Effect of Background Noise on Food Perception. Food Quality and Preference. 22(1). pp 42-47.

Webster, J. C., & Klumpp, R. G. (1962). Effects of Ambient Noise and Nearby Talkers on Face-to-Face Communication Task. Journal of the Acoustic Society of America. 34(7). pp. 936–941.

Williams, F. (1968) Reasoning With Statistics. Holt, Rinehart and Winston. New York.

Zampini, M. and Spence, C. (2010) Assessing the Role of Sound in the Perception of Food and Drink. Chemosensory Perception. 3(1). pp 57-67.

Zwicker, E. & Fastl, H. (1999) Psychoacoustics - Facts and Models. Springer Verlag, Berlin.

Figures

Literaturcafe (2013) Abtauchen in neue Hörwelten. http://www.literaturcafe.de/html/buecher/verne/ophp/ Accessed 2013-07-02.

Wikipedia. (2013) A-weighting. http://en.wikipedia.org/wiki/A-weighting Accessed 2013-07-02.

Appendix

RESTAURANTS

Cafés

1. Espresso House, Regeringsgatan 52
2. Ritorno, Odengatan 80
3. Espresso House, Odengatan 43
4. Östra stationscafét, Valhallavägen 77
5. Stories KTH, Osquars Backe 25
6. Linquists konditori, Odengatan 27
7. Franska Caféet NK, Hamngatan 18
8. Café Panorama, Kulturhuset Sergels Torg 3
9. Brydlings Café, Åhléns, Klarabergsgatan 50
10. Kungstornet, Kungsgatan 33

Fastfood Restaurants

1. Burger King, Odenplan
2. Subway, Odengatan49
3. Kebabkungen, Odengatan 54
4. McDonalds, Sveavägen 71
5. Kungshallen, Kungsgatan 44
6. McDonalds, Hamngatan 8
7. MAX, Kungsträdgårdsgatan 20
8. McDonalds, Regeringsgatan
9. Burger King, Sergels Torg 1
10. McDonalds, Nybrogatan 18
11. Hisar Fastfood, Hötorgshallen, Hötorgshallen 27

Lunch Restaurants

1. KTH Nymble, Drottning Kristinas väg 15
2. Q-restaurangen KTH, Osquldas väg 4
3. Vigårda, Norrlandsgatan 13
4. Teaterbaren Kulturhuset, Sergels Torg 3
5. Hurry Curry, Slöjdgatan 11

6. NK Lunch, Hamngatan 18
7. Brydlings Åhléns , Klarabergsgatan 50
8. Brasserie Tures, Hamngatan 37
9. Restaurang Borggården, Armémuseum, Artellerigatan 13
10. Vurma, Vasastan Gästrikegatan 2

On the next page: the questionnaire used as inspiration for collection of on-site annotations and rating of perceptual features in each restaurant.

SURVEY ON THE SONIC QUALITY OF RESTAURANTS

This survey is part of a research project carried out on the Department of Speech Music and Hearing, KTH. The project aims to investigate the sonic quality of various restaurants in Stockholm. For questions and queries, please feel free to email emmafrid@kth.se or permagnus@ntu.edu.sg.

YOUR AGE: _____ **YOUR GENDER:** □ Female □ Male □ Other

1. **HOW MUCH TIME ARE YOU PLANNING TO SPEND IN THIS RESTAURANT?**
   - □ 15 min
   - □ 15 – 30 min
   - □ 31 – 60 min
   - □ > 60 min

2. **WHAT DO YOU THINK OF THE LEVEL OF BACKGROUND MUSIC IN THIS RESTAURANT?**
   - □ Can't hear any music
   - □ Too soft
   - □ It's OK
   - □ Too loud

3. **RATE THE LEVEL OF AMBIENT NOISE (1-5) IN THIS RESTAURANT: (circle your choice: 1 = no noise … 5 = extremely loud noise)**

   1 2 3 4 5

4. **RATE HOW THE SOUNDS IN THIS RESTAURANTS AFFECT YOUR OVERALL DINING EXPERIENCE: (circle your choice: 1 = enhances greatly … 5 = ruins completely)**

   1 2 3 4 5

5. **DESCRIBE 3 CHARACTERISTIC SOUNDS OF THIS RESTAURANT. DO YOU LIKE THESE SOUNDS (circle YES/NO):**

   Sound 1 : ____________________________________ YES NO
   Sound 2 : ____________________________________ YES NO
   Sound 3 : ____________________________________ YES NO

6. **WHAT IS YOUR OVERALL SONIC IMPRESSION OF THIS RESTAURANT? (circle your choice)**

   Very good Good Neutral Bad Very bad

7. **OVERALL, HOW DOES THIS RESTAURANT MAKE YOU FEEL?**

   ____
   ____

please turn over…

SURVEY ON THE SONIC QUALITY OF RESTAURANTS

This survey is part of a research project carried out on the Department of Speech Music and Hearing, KTH. The project aims to investigate the sonic quality of various restaurants in Stockholm. For questions and queries, please feel free to email emmafrid@kth.se or permagnus@ntu.edu.sg.

YOUR AGE: _____ **YOUR GENDER:** □ Female □ Male □ Other

1. **HOW MUCH TIME ARE YOU PLANNING TO SPEND IN THIS RESTAURANT?**
   - □ 15 min
   - □ 15 – 30 min
   - □ 31 – 60 min
   - □ > 60 min

2. **WHAT DO YOU THINK OF THE LEVEL OF BACKGROUND MUSIC IN THIS RESTAURANT?**
   - □ Can't hear any music
   - □ Too soft
   - □ It's OK
   - □ Too loud

3. **RATE THE LEVEL OF AMBIENT NOISE (1-5) IN THIS RESTAURANT: (circle your choice: 1 = no noise … 5 = extremely loud noise)**

   1 2 3 4 5

4. **RATE HOW THE SOUNDS IN THIS RESTAURANTS AFFECT YOUR OVERALL DINING EXPERIENCE: (circle your choice: 1 = enhances greatly … 5 = ruins completely)**

   1 2 3 4 5

5. **DESCRIBE 3 CHARACTERISTIC SOUNDS OF THIS RESTAURANT. DO YOU LIKE THESE SOUNDS (circle YES/NO):**

   Sound 1 : ____________________________________ YES NO
   Sound 2 : ____________________________________ YES NO
   Sound 3 : ____________________________________ YES NO

6. **WHAT IS YOUR OVERALL SONIC IMPRESSION OF THIS RESTAURANT? (circle your choice)**

   Very good Good Neutral Bad Very bad

7. **OVERALL, HOW DOES THIS RESTAURANT MAKE YOU FEEL?**

   ____
   ____

please turn over…

8. TO WHAT EXTENT DO YOU AGREE WITH THE 8 STATEMENTS BELOW ON HOW YOU EXPERIENCE THE PRESENT SURROUNDING SOUND ENVIRONMENTS? (please tick off one response alternative per statement)

The sound environment is:	Strongly agree	Agree	Neutral	Disagree	Strongly disagree
Pleasant	☐	☐	☐	☐	☐
Chaotic	☐	☐	☐	☐	☐
Exciting	☐	☐	☐	☐	☐
Uneventful	☐	☐	☐	☐	☐
Calm	☐	☐	☐	☐	☐
Annoying	☐	☐	☐	☐	☐
Eventful	☐	☐	☐	☐	☐
Monotonous	☐	☐	☐	☐	☐

9. OVERALL, HOW WOULD YOU RATE THE VISUAL ENVIRONMENT IN THIS RESTAURANT?

Very good	Good	Neutral	Bad	Very bad
☐	☐	☐	☐	☐

8. TO WHAT EXTENT DO YOU AGREE WITH THE 8 STATEMENTS BELOW ON HOW YOU EXPERIENCE THE PRESENT SURROUNDING SOUND ENVIRONMENTS? (please tick off one response alternative per statement)

The sound environment is:	Strongly agree	Agree	Neutral	Disagree	Strongly disagree
Pleasant	☐	☐	☐	☐	☐
Chaotic	☐	☐	☐	☐	☐
Exciting	☐	☐	☐	☐	☐
Uneventful	☐	☐	☐	☐	☐
Calm	☐	☐	☐	☐	☐
Annoying	☐	☐	☐	☐	☐
Eventful	☐	☐	☐	☐	☐
Monotonous	☐	☐	☐	☐	☐

9. OVERALL, HOW WOULD YOU RATE THE VISUAL ENVIRONMENT IN THIS RESTAURANT?

Very good	Good	Neutral	Bad	Very bad
☐	☐	☐	☐	☐

Axelsson, Ö., Nilsson, M.E., and Berglund, G. (2011) Swedish Soundscape Quality Protocol (SSQP, March 2011).

Figure 5.1. Example of typical café environment.

Figure 5.2. Example of typical café environment.

Figure 5.3. Example of typical restaurant environment.

Figure 5.4. Example of typical restaurant environment.

Figure 5.5. Example of typical fastfood environment.

Figure 5.6. Example of typical fastfood environment.

Restaurant	Materials and architecture	Characteristic sounds
c1	stone floor multiple shops and escalators in the vicinity open space in the center of the room	chairs scraping against floor cutlery speakers voice (commercial)
c2	metallic tables stone floor stone brick walls large glass wall facade	waitress talking door opening conversations chairs scraping against floor
c3	situated at the top of the building wooden floor wooden tables glass walls towards Sergels Torg	cutlery laughter conversations children
c4	wall-to-wall carpet some sofas wooden tables low ceiling antique furniture wooden tables	fan children door opening and closing
c5	tile floor large glass windows very small café	kitchen sounds passing commuter train passengers coffee machine milk foamer
c6	soft sofas plastic floor low ceiling due to the escalator metal surfaces around escalator large glass wall	escalator coffee machine milk foamer
c7	soft cushions wooden floor some curtains low ceiling large windows	espresso machine coffee machine traffic
c8	tile floor soft chairs antique style but newly renovated long curtains covering main part of the windows	humming fridge children conversations
c9	sofas no curtains wooden floors soft chairs wooden table pretty small room	background music conversations sounds of plates and dishes
c10	50s design wood soft wall-to-wall carpet leather sofas wooden tables tile floor	conversations background music traffic coffee machine
f1	designed space no curtains stone floor only hard chairs	kitchen sounds fridges microwave beeps background music sounds from the till (coins)
f2	tile floor wooden tables	background music traffic kebab meat cutter
f3	low ceiling glass walls (fake) leather sofas stone floor	conversations background music microwave beeps from kitchen
f4	tile floor typical american subway interior design (plastic)	fridge traffic sounds from the till (coins) sound from the till
f5	circular room multiple restaurants renovation upstairs linoleum floor hole in roof escalator	escalator conversations kebab meat cutter dish racks moving
f6	not many seats upstairs several glass walls low roof hard surfaces	microwave beeps traffic sounds from the till (coins)
f7	glass walls stone floor open areas small tables and small chairs	kitchen sounds rustling papers chairs scraping against floor
f8	tile floor hard chairs plastic tables no textiles or decor	background music rustling papers conversations
f9	linoleum floor circular glass wall facing Sergels Torg plastic tables no textile some leather sofas	background music sounds from Sergels Torg conversations
f10	tile floor tile on some walls open space	microwave beeps from the kitchen machines background music chairs scraping against floors
f11	tiles on all walls wooden tables plastic floors small space few chairs within food market hall	kebab meat cutter humming fridge telephone microwave beeps
lr 1	curtains large windows wooden floor some bean bags high ceiling	waiter picking salt and pepper shakers from tables conversations footsteps
lr 2	many cushions small cafe well furnished	conversations background music coffee machine
lr 3	two floors large glass windows tile floor wooden furniture no textiles wooden floor upstairs area	background music trays scraped exclamations of orders
lr 4	wooden floor high ceiling large glass facade	screaming babies cutlery background music
lr 5	wooden floor a lot of decor wooden chairs no curtains almost no textile	dish room background music glasses
lr 6	carpets on 50 percent of the floor relatively small windows stone floor decor in roof	children chatter cutlery
lr 7	wooden floor wooden tables large space two different tills	sounds from escalator coffee machine milk foamer
lr 8	textile lamps very high ceiling soft chairs glass facade and stone floor	people conversations scraping sounds children dishes
lr 9	stone roof and floors vault wooden chairs and tables no textile	cutlery against plates chairs scraping against floor conversations
lr 10	large glass windows wooden floor curtains on the majority of the walls some sofas	plates and cutlery laugher conversations child

Table 5.1. Overview of collected data used for word frequency analysis.