NTNU
Norwegian University of
Science and Technology

Method for classification of restaurant acoustics

Thomas Wøhni

Master of Science in Electronics
Submission date: June 2018
Supervisor: Odd Kr. Pettersen, IES
Co-supervisor: Olav Kvaløy, Sintef

Norwegian University of Science and Technology
Department of Electronic Systems

Summary

This master's thesis presents a method for classification of the acoustic quality of restaurants. Noise in restaurants makes verbal communication difficult and lowers the quality of the dining experience. It is a common problem that most people can relate to. Still, little is done to improve the situation. The goal of the method presented here is to raise awareness in society on this topic. With this goal in mind, the focus of the method is on simplicity and ease of implementation. The method is based on a single measurement of ambient noise level (both dBA and 1/3 octave spectrum), in occupied state. It requires no active involvement from the restaurants such as access during unoccupied periods. This will hopefully enable a large portion of restaurants to be assessed, over time.

One complicating factor in restaurants is the feedback relationship between speech level and background noise. People tend to speak louder when subjected to noise (the Lombard effect). In multitalker situations where the ambient noise is mostly other talkers, the result can be a rapid escalation of noise. This needs to be factored in when restaurants are evaluated, especially if a single measurement is used. Considering that restaurants are usually never completely full, the maximum noise level is based on 80% occupancy. Since it's difficult to predict when any restaurant is at 80% occupancy, the method must be effective for lower rates. In such cases, the maximum ambient noise level is estimated, factoring in the Lombard effect. This estimation process is complicated by cross table distance and the method includes a step for factoring that in.

The measured (or estimated from measured) ambient noise level at 80% is compared to the speech level. Speech level is assumed from the available models of the Lombard effect, based on standardized statistical models. The resulting ratio between speech and noise (SNR) is used for classification of the acoustic quality. The classes are based on ISO 9921:2003, the international standard for assessment of speech communication. The 1/3-octave spectrum measurement is used for calculation of the SII (Speech Intelligibility Index). The SII is used here to support the conclusion drawn from the dBA measurement.

The end result is presented in three simple terms; ”Good”, ”Okay” and ”Bad”, with corresponding smiley-face icons. The purpose is to communicate to the potential guests what kind of experience they can expect. This will hopefully raise awareness and provide context for their subjective experiences. The dividing line between good and okay is set to 59 dBA (ambient noise level). I.e the measured or estimated maximum noise level is 59 dBA or less, giving an SNR of 3 or higher. The line between ”okay” and ”bad” is set to 71 dBA, for an SNR of -3 or worse. The classes are defined to give a realistic view of the kind of experience guests will have. Most likely, very few restaurants will achieve the ”good” rating at present time. This is unfortunately an accurate description of the situation and the motivation for the method itself.

Preface

With my background from the world of music, the choice to study acoustics was easy. I've recorded sound since I was barely a teenager and worked in recording studios ever since. I've always known that my love of music and sound technology was slightly above the norm. However, it did surprise me to learn how little the engineering world cares about acoustics. I've had many conversations with acousticians who've said similar things: The acoustics are *not* a priority when buildings are constructed. Many acoustical problems could have been avoided *at no extra cost* if the acousticians had been consulted earlier in the process. Also, that it's just about building something to an "acceptable standard", and preferably not better (as this surely must cost too much!).

With this in mind, the chance to focus on restaurant acoustics immediately attracted me. This is something the general public can relate to. If consumers became more aware of noise problems, the business world would have to respond. In restaurants, we've all experienced that noisy atmosphere and felt the joy of conversation quickly fade. People who are sensitive to noise often avoid dining out completely. And yet restaurants keep getting away with their "noise crimes" . The only explanation I can think of is, people believe it's unavoidable.

The topic of restaurant acoustics is also fascinating for many other reasons. It's a field that combines the physics of sound, the psychology of social behavior with the physiology of human hearing. And after the acoustician is done evaluating, the results must be communicated to the restaurant manager. So there is an element of business in the mix.

Finally, I would like to thank my supervisor, Olav Kvaløy. His knowledge about acoustics and scientific writing was very helpful.

Table of Contents

List of Tables

List of Figures

Abbreviations

EE = Eating establishment(s)
AC = Acoustic capacity
$L_{N,A}$ = A-weighted ambient noise level
$L_{S,A}$ = A-weighted speech level, 1 meter from the mouth
T = reverberation time in seconds
V = room volume in m^3
A = Absorption area in metric Sabins

Chapter 1

Introduction

Going to a restaurant is about having an *experience*, not just to eat. It's the conversation, with old and new friends. To sit down for a meal is a social event where people communicate. Ideally, the transmitted messages are received and understood without too much effort. This is the field of acousticians concerned with *speech intelligibility*. It's an important part of every building intended for humans who talk and listen. Despite this, reviews of restaurants often contain little discussion of the acoustics. Both professional food journals and web-based customer reviews often skip it entirely. This pattern is evident in people's dining habits as well. One study found that noise effects ”satisfaction, but not loyalty ” (Raab et al., 2013). Suggesting that people will complain, but not change their habits based on noise levels. This raises an interesting question. If ”everybody” has felt the problem at some point, why is it allowed to go on? One possible explanation is people assume the noise is unavoidable. ”Something we just have to live with” . The goal of this work is to challenge this perception and raise the general awareness of acoustics.

At the beginning of writing this thesis, about 15 restaurants were contacted via email. The goal was to get an overview of the general situation. The content of the email can be summarized in one sentence. Free acoustic advice at no risk or cost to the restaurant. Of the 15, only 3 replied. After a quick preliminary assessment of these establishments a few things became clear. For one, the restaurant managers are aware of the noise problem. This could certainly explain the low response rate. Secondly, restaurants often have some (but too little) acoustic treatment. Typically absorptive materials, installed in the ceiling. This insufficient attempt at improving the acoustic quality proved to be a warning sign. Because more importantly, they have no real intention of improving the space further. No real desire to invest and raise the quality of the dining experience, for their guests. Although this overview was not obtained in any scientific manner, it did make me rethink the strategy. Especially after consulting with more experienced acousticians who agreed with these ”findings” .

Some have even said, part of the job is convincing people acoustics actually matter. Unfortunately for acousticians, restaurants don't go out of business just because of bad acoustics. If fact, the noise problems in restaurants are often due to many people speaking simultaneously. To a restaurant owner, this is great news. To make the case that a restaurant desperately needs acoustic treatment, while business is ”booming”, is difficult. There are no scientific proofs that fixing the acoustics will automatically lead to higher profits. How can noise problems in restaurants ever be solved, if the restaurant sector doesn’t want to prioritize it?

In this work, the choice became to focus on the guests. The idea is simple: Create a method for restaurant assessment that is *simple* to implement, provides useful information to potential guests and communicates it in everyday terminology. A simple method has a better chance of actually being used on a large scale. Useful information would mean giving the guests some validation for their experience (”yes, this *is* too loud”). Communicating the result in an uncomplicated way means no mention of Hertz or Helmholtz. Why say it in complicated terminology when you can say it with a smiley face emoticon? This will hopefully accomplish two things. First, it will provide reviewers and guests some context for their experience. It sends a signal that loud restaurants are *not* unavoidable. Secondly, this raised awareness will hopefully result in clear market trends. Over time, and with enough restaurants evaluated, the free market will decide. If noise levels truly matter to people, restaurants with bad acoustics *will* suffer. Guests may start to notice loud restaurants more and associate them with ”bad acoustics”, not ”this is unavoidable” . Eventually, they will seek to know the acoustic quality *before* they chose where to eat.

A simple and practical method also has a better chance of being implemented on a large scale. In order to effect a lasting change, a large portion of restaurants needs to be assessed. It would not be of much help to ”attack” a few, if the majority are left to carry on as usual. There are ethical concerns to consider, as well. A negative assessment of a particular restaurant could have consequences for their survival. Evaluating a small and randomly selected group and ”sounding the alarm” could ruin them. In the spirit of free markets, let the audience decide. This thesis aims to give them the information they need to do so.

Chapter 2

Literature Review

Four main questions needs exploring: (1) What are the metrics that defines acoustic quality in a restaurant? (2) Which factors will have an influence these metrics? (3) How can this be measured and evaluated? (4) How can the evaluation be simplified and communicated in a way that is both helpful and educational to the public?

There is a lot of relevant research available, dating all the way back to 1911, with the discovery of the Lombard effect. However, most (if not all) of it has a different goal than the method proposed here. The existing research is focused more on *prediction*, not on simple and practical evaluation methods. This chapter is a brief outline of the existing work and a discussion on how it can be used in this context.

The measure of acoustic quality in a venue depends on what the intended use is. A concert hall should be judged on different criteria than a conference room. Due to the social nature of restaurants, *verbal communication* is the primary acoustical concern. The ease and quality of verbal communication can be measured by comparing the speech level to the ambient noise level, in various ways. In figure 2.1 below, an overview is shown.

Several authors have contended that *speech intelligibility* is a suitable measure of a acoustic quality quality(van Heusden et al. (1979), Bradley (1986), Lazarus (1987) etc). Many international standards exist, relevant to this research topic. ISO 9921 (on assessment of speech communication) provides a guideline for noise levels in various situations. It is based on the idea of comparing speech levels and noise level to determine intelligibility. ANSI 3.5-1997 quantifies the *Speech Intelligibility Index* (SII) and provides an objective measure of speech intelligibility, from speech spectrum level and noise spectrum level. This leads to a need for predicting both speech and noise spectrum levels.

Interestingly, in a restaurant these levels are *linked*. Most of the background noise in a restaurant is other people talking. And, the talkers' vocal output power is not constant, it increases as the noise increases. This creates a feedback system where the noise can quickly get out of hand. Since the 1950s, this phenomenon has been informally referred to as the ”Cocktail party effect” .

In 1911, Etienne Lombard discovered that people increase their speech levels involuntarily when subjected to noise. Over the years, many authors have tried to quantify

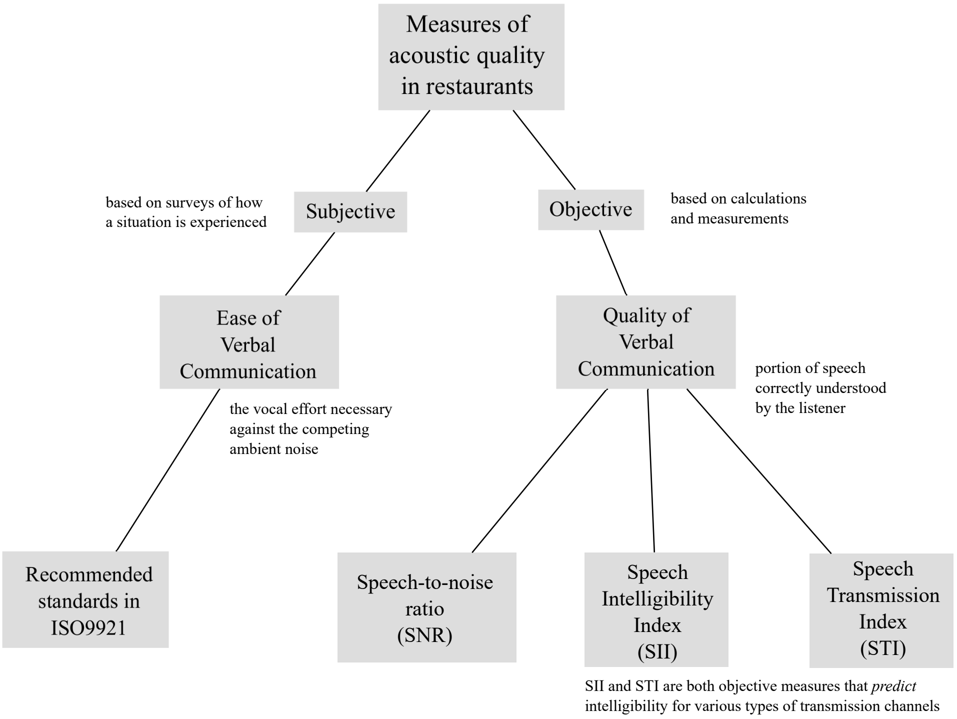


Figure 2.1: The various ways to measure the acoustic quality of a restaurant

the effect (Korn (1954); Webster and Klumpp (1962); van Heusden et al. (1979); Lazarus (1986);). The main problem has been to establish how big the increase in vocal effort is, for each dB increase in noise. This value is called the *Lombard slope*. Korn (1954) found that the effect starts at noise levels above 45 dBA and suggested a Lombard slope of 0.38 for noise levels above 55 dBA. Pickett (1958) used an anaechoic chamber and examined the Lombard effect in a free field situation. The results suggested a Lombard slope of 1. Webster and Klumpp (1962) used various types of noise and found a slope of 0.5 for ”thermal noise” and a slope of 1 for vocal noise. They also commented on the different results in other studies and suggested that perhaps ”the requirement for accuracy” in a given situation could influence the speaker’s output level. Lombard slopes ranging from 0.22 to 1 was reported throughout the following decades and although interesting, the research is regarded as statistically weak (Jr. et al., 1989). The ISO 9921 standard from 2003 also takes the Lombard effect into account, but does not make it clear what slope it uses.

In recent times, several models of varying complexity have been introduced. Hodgson et al. (2007) suggested that the effect is *non-linear*, and less pronounced at both low and high vocal levels. He also found large differences between different individuals and different situations. This could explain the varying results in earlier studies. He arrived at the non-linear model for the Lombard slope using optimization techniques on measured data.

The Lombard effect becomes most influential in situations where the background noise *is* mainly other speakers. A simplified model for predicting noise levels (dBA) in restaurants was proposed by Rindel (2010), *Acoustic Capacity* (AC). It is an attempt at communicating an EEs acoustic limitations in a single number rating. The AC model includes room acoustic parameters such as room volume, reverberation time and equivalent absorption area. The output is the acoustic capacity of the room, i.e. the maximum number of guests before verbal communication becomes difficult. Rindel uses 45 dBA as the starting point of the Lombard effect, as established by Korn (1954) and choses a slope of $c = 0.5$. This is in contrast to the non-linear model proposed by Hodgson et al. (2007). Instead, Rindel chooses the mid-point of non-linear model as his linear slope. Being a statistical model, it is only valid for groups larger than 50 people. This is a drawback for a real-world method where many EEs are smaller.

Customer satisfaction is a subject of great importance in the corporate world. Any product or service offered to the public is shaped by research. Both what the product or service *is* and how it's marketed. The *servicescape*, i.e the physical environment in which a service is delivered, is a popular umbrella-term for various sectors (Bitner, 1992). A restaurant falls into the subcategory of a *hospitality environment*. Many publications exists, such as The Journal of Foodservice Business Research, the Journal of Culinary Science & Technology and the International Journal of Hospitality & Tourism Administration. They publish research work done on customers responses and behavioral patterns, usually described as hospitality research. In recent years, this field has started to focus more and more on the effects of unwanted ambient noise (Zemke et al., 2011). The method presented here is intended to define the acoustic comfort level in a given restaurant, from a simple measurement. So, linking the available hospitality research to the acoustic research field is a key topic.
The first point of interest would be *demographics*. People respond differently based on age, education level, socio-economic level and so forth. Some of the hospitality research contains results that can be of value for the acoustician. One study, conducted in a restaurant in the southwestern United States, found *age* to be the most significant factor (Zemke et al., 2011) in noise perception. This is perhaps no surprise, as hearing ability diminishes with age. The other main factor was education level. People with higher education are more discerning and critical of noise pollution in restaurants. Across genders the only significant difference was that women generally experienced the ambient noise as less intrusive than the men.
Several studies have shown that noise from other guests is more intrusive than mechanical noises from the surroundings ((Raab et al., 2013), (Novak et al., 2010), Zemke et al., 2011). If guests can understand the words being said at nearby tables, it is even more distracting. This points at the complexity of the whole topic. There is a psychological component to the perceived nuisance.
A study of 112 restaurants in Singapore (Lindborg, 2015) found that price level and noise level was negatively correlated. The more expensive restaurants had less ambient noise. This study was performed by acousticians and focused on the materials used in the interior design. A possible explanation for this is that when people spend more money they become more demanding as customers.

A Chinese study the acoustic comfort of large dining spaces (Chen and Kang, 2017) found that people who dine out more frequently are less sensitive to noise than those who go less often to restaurants. Put simply, noise matters for the sector as a whole, but not (yet) for individual restaurants. People who've had negative noise experiences simply go out less often. There is basically an untapped market of noise sensitive people out there. And yet, the restaurant industry, which is struggling in many parts of the world, is still not responding.

In summation, the existing work on Lombard and multitalker situations has been focused more on prediction of noise levels and design of rooms. This thesis will use these results to create a less complex evaluation method, with the hope that it might raise awareness in society.

Chapter 3

Theory

In this chapter, the theory needed for assessment of a restaurant is described. Starting with how sound propagates and the foundation of room acoustics, the discoveries of Wallace Sabine. Next, the tendency of humans to raise their vocal effort when speaking in noise is discussed. This effect was discovered more than a hundred years ago, and like many things relating to human cognitive process and hearing, it is still not entirely understood or quantified.

In recent years, the work of Sabine and Lombard has been combined into *Acoustic Capacity* (Rindel, 2010). Acoustic capacity refers to the maximum number of people a room can hold, before the noise level will reach unacceptable levels. In this context, *unacceptable* refers to noise levels for which verbal communication becomes difficult.

The simplest objective measure of verbal communication is the ratio between speech level and ambient noise level. The accepted standards for verbal communication is described in a separate section. These recommended levels are also used as a baseline in the Acoustic Capacity concept.

3.1 Definitions

3.1.1 Ambient noise

The word *noise* can mean different things depending on context. Generally it's thought of as ”unwanted sound” and a form of pollution. In a restaurant however, there can be other sounds, such as music playing in the background. A natural question would be if this sound should be separated from the unwanted noise. And this may be an interesting avenue for future work, but in the context of assessing a restaurant as *it is now*, it can be argued that it does not matter. If a restaurant has ambient noise level above the recommended limits, the problem is mroe likely to be too long reverberation time and/or too many people speaking simultaneously.

So, in this thesis the term *noise* will be used to describe all sound competing with the speaker's voice. In other words, the *ambient noise level* (with symbol $L_{N,A}$) is the A-

weighted level of the diffuse sound field in the room. This noise is what is disturbing the conversation and causing problems for both the speaker and the listener.
The ratio between speech level and noise level is a common measure of the acoustic quality in a restaurant. The common term for a ratio bewteen wanted and unwanted sound is *signal-to-noise ratio*, or SNR. In the context of assessing a restaurant's acoustic quality, the *signal* is the speech of a guest and the *noise* is the ambient noise level in the room. The term *speech-to-noise ratio* is commonly used interchangeably with *signal-to-noise ratio*. In the following, SNR will be used in the same way; the ratio between the A-weighted speech level ($L_{S,A}$) measured 1 meter from the mouth of the speaker and the A-weighted ambient noise level at the listener's ear ($L_{N,A}$).

3.1.2 The Cocktail party effect vs the Lombard effect

The informal term ”Cocktail party effect” can have two different meanings. It can either refer to how people raise their vocal effort when speaking in noise, or it can refer to the brain's ability to focus on one particular voice in a noisy environment. The tendency to speak louder when subjected to noise is more formally named the ”Lombard effect”, named after the man who discovered it. This term is what will be used in this thesis. The other phenomenon of the minds ability to pick out speech from noise is also relevant in restaurants, and will be discussed on it's own. To avoid unnecessary confusion it will not be referred to as the cocktail party effect.

3.2 Room acoustic parameters

The foundation for a discussion involving room acoustics, such as in restaurants, will be the discoveries of Wallace Sabine (1868-1919). His work can be found in every acoustics textbook to this day, such as Kinsler et al. (2000) or Long (2005) . Sabine arrived at an empirical relation between the reverberation time of a room, its size and the amount of absorbing material present. Sabine's equation

$$T \propto \frac{V}{A} \tag{3.1}$$

relates the *room volume* V, *total absorption* A and *reverberation time* T. The underlying assumption is that sound radiates as *rays*, traveling outward from the source. The rays will travel until they hit a surface and be reflected back. After a large number of reflections the sound is assumed to be *diffuse*. In a *diffuse* sound field, the energy density is the same throughout the field. In other words, the sound pressure level is the same throughout the space. This is a simplification of how sound works in a room, as it neglects phenomenons such as *standing waves*. It also requires a large number of reflections before the sound truly can be called diffuse. Even so, with proper values chosen for A, Sabine's equation leads to valid conclusions.

3.2.1 Reverberation time

In restaurants, one of the most common noise complaints is the sound of other people talking. If one wants to decrease this noise level, the first parameter to look at is reverberation time. The acoustic energy in the room will bounce around and be reflected from any hard surface. When the sound source is stopped, the reflected sound from the enclosure will keep bouncing for a little time period. The *reverberation time* T is defined as the time (in seconds) it takes for the level of the sound to drop by 60 dB. Assuming a sound speed of $c = 343\frac{m}{s}$, this can be expressed in the famous *Sabines reverberation formula*:

$$T = 0.161\frac{V}{A}\ [s] \tag{3.2}$$

3.2.2 Absorption

The absorptive materials in a room are materials that will absorb some of the acoustic energy and transform it into other forms, such as heat (through friction). The quantity A is the *total sound absorption*, also called *absorption area*. It is a sum of all the individual absorptive surfaces S_i in the room multiplied by their absorption coefficient:

$$A = \sum_i A_i = \sum_i S_i a_i\ [sabins] \tag{3.3}$$

where a_i is the *Sabine absorbtivity* for the ith surface of area S_i.

A simpler way to express this is to think of it as the area of an open window, that reflects no sound back. If all the aborbtive materials in the room where removed and a window installed, with an absorbtion coefficient of 1 (no reflection), what would the area of that window need to be, if the room was to have the same reverberation time.

Absorption can also be expressed by an average coefficient, for all the surfaces. If the total surface area in the room is S, the *average Sabine absorbtivity* $\overline{a}$ is defined by

$$\overline{a} = \frac{A}{S} \tag{3.4}$$

Generally speaking, of the acoustic variables in a room, absorption is the simplest to change. If the desire is to shorten the reverberation time, adding absorptive materials will improve the situation. The exact area of absorption that is needed depends on the material itself and the desired T and the spectral content of the sound. In restaurants, most of the energy of the ambient noise is concentrated in the 200-500 Hz range. This is where the fundamental frequencies of adult voices are. Another quantity related to the absorption is the *critical distance*. The critical distance of a room is the distance from the source at which the direct sound and the reverberant sound has equal amplitude.

3.2.3 Diffuse field theory

When a sound source starts producing sound in a room, reflections at the walls will produce a sound energy distribution that becomes more and more uniform over time (Kinsler et al., 2000). The sound in the room will become more and more diffuse and, in theory, be distributed equally. This is the basic assumption of diffuse-field theory. The sound field is diffuse if the following two conditions apply: (1) at any position in the room the reverberant sound waves are incident from all directions with equal intensity (and random phase relations); (2) the reverberant sound field is the same at every position in the room.

However, this is not necessarily valid in every room. Exceptions can be rooms with deep recesses, rooms with sound focusing surfaces or rooms that are coupled with other rooms. All these cases can occur in restaurants, as they are not always simple rectangular shapes. The question of when diffuse field theory is difficult but some general guidelines exists (Hodgson, 1996). Hodgson devoted an article to the subject and wrote:

> " Diffuse-field theory is used by practitioners to predict sound fields in rooms of every type. Often forgotten is the fact that the theory is based on assumptions which may limit its applicability. "

Hodgson considered four main parameters that will play a role in the accuracy of diffuse field theory. (1) Room shape, described by the aspect ratio (length:width:height). (2) Surface absorption, described by how it's placed spatially and it's magnitude described by the average surface absorption coefficient. (3) Surface reflection, as described by the diffuse-reflection coefficient, equal to the proportion of reflected energy which is reflected diffusely. (4) fittings (or volume scatterers)- these are obstacles in the room volume that scatter sound randomly, as described by their average scattering cross-section volume density (the average effective sound scattering area divided by the volume of that part of the room being considered as fitted). In restaurants all four parameters might have values that complicate matters. It's absolutely certain that there will be uneven surfaces that scatter sound in random ways. There might also be obstacles to the sound (bar section, concrete pillars etc) that will alter the nature of the reflected sound. However, since the method proposed here is not aimed directly at predicting reverberation time, the concerns about reflecting surfaces and scattering volumes will be put aside. It is however an area that should be explored in further work, to test the method.

Another significant factor is the frequency of the sound. As frequency increases, the natural resonances in the room will be spaced closer in frequency. The result is more interference between standing waves and thus a more random sound field. The frequency above which a room can be assumed to have fairly diffuse sound field is called the *Schroeder frequency*. The Schroeder frequency is given by:

$$f \geq 2000\sqrt{\frac{T}{V}} \tag{3.5}$$

where T is reverberation time in seconds and V is the room volume.

In restaurants, the main sound source (people talking) will have most of it's energy in the 200-500 Hz range. This is above the Schroeder frequency in any realistic restaurant scenario so this will not hinder the diffuse field assumption.

3.3 Estimating speech level

In order to determine the quality of verbal communication from SNR, the speech level must be known. For practical reasons, the only realistic option is to statistical data, determined in complicated controlled tests and standardized for uses like this. The increase in vocal effort when speaking in noise is described in ISO 9921 - *assessment of speech communication*. It also provides recommended levels of speech-communication quality, in different situations. For restaurant acoustics, the relevant part of the standard is ”prolonged normal person-to-person communication” . For this situation ISO 9921 recommends a *Minimum intelligibility rating* of ”Good” and a *Maximum vocal effort* of ”Normal”. The details are described in section 5.3 of the standard.

3.3.1 The Lombard effect

The Lombard effect describes how people raise their vocal effort when subjected to noise. When background noise increases, people compensate by raising their voice. The increased vocal effort results not only raises speech levels, it also changes the spectral content. As shown in figure 3.7 on page 21, the energy is focused slightly higher in frequency, especially at very loud levels. It is still not known how much of the increase in vocal intensity is due to an automatic reflex (Lombard effect) or due to a learned response to listener’s needs (Jr. et al., 1989).
In situations where the ambient noise consists mainly of people talking in a room, the Lombard effect creates a feedback system. Since the sources (people speaking) are not constant power, there is an unusually strong effect from increasing the number of sources. As illustrated in figure 3.1, doubling in the number of guests result in a 6 dB increase in ambient noise level, not the 3 dB one would expect if the sources had constant power.

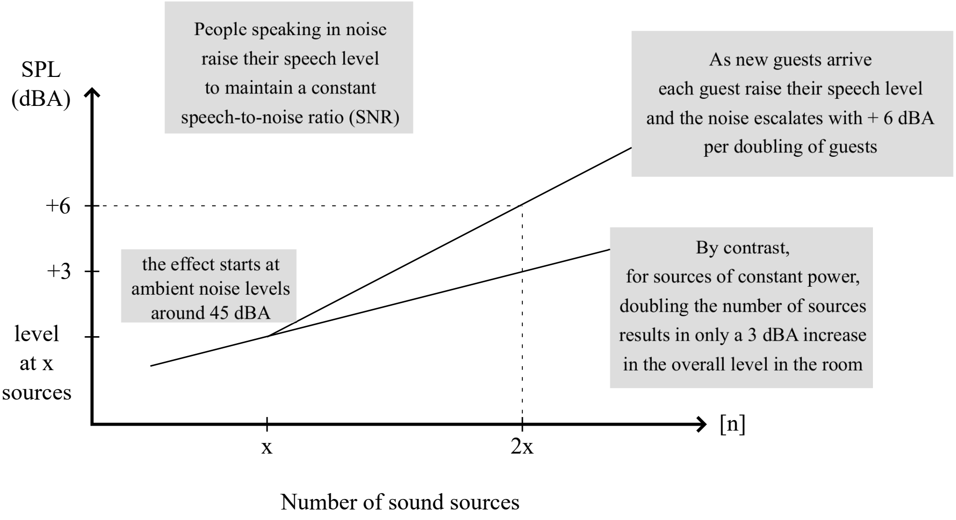


Figure 3.1: The variable output of people speaking in noise

Linear Lombard model

The assumption made in the previous section is that the effect is *linear*. That is, to assume a linear relationship between the increase in vocal effort and ambient noise level. I.e. the speech level increases at the same rate, through the vocal range of the talker. The linear model is used for simplicity and because the data available on the Lombard effect is considered statistically weak (Rindel, 2010). The effect has been established to start at noise levels of 45 dBA (Korn, 1954) and voice levels of 55 dBA. The increase of speech level as a function of the A-weighted ambient noise level is described by the slope *c*, the *Lombard slope*. Equation 3.6 expresses the relationship:

$$L_{S,A,1m} = 55 + c \cdot (LN, A - 45)\, , (dB) \tag{3.6}$$

where LN, A is the ambient noise level and $L_{S,A,1m}$ is the predicted A-weighted speech level, at 1 m distance. For context, ISO 9921 defines a speaking level of 54 dBA is as ”Relaxed” .
The relationship between speech level and ambient noise is plotted in Figure 3.2 . This

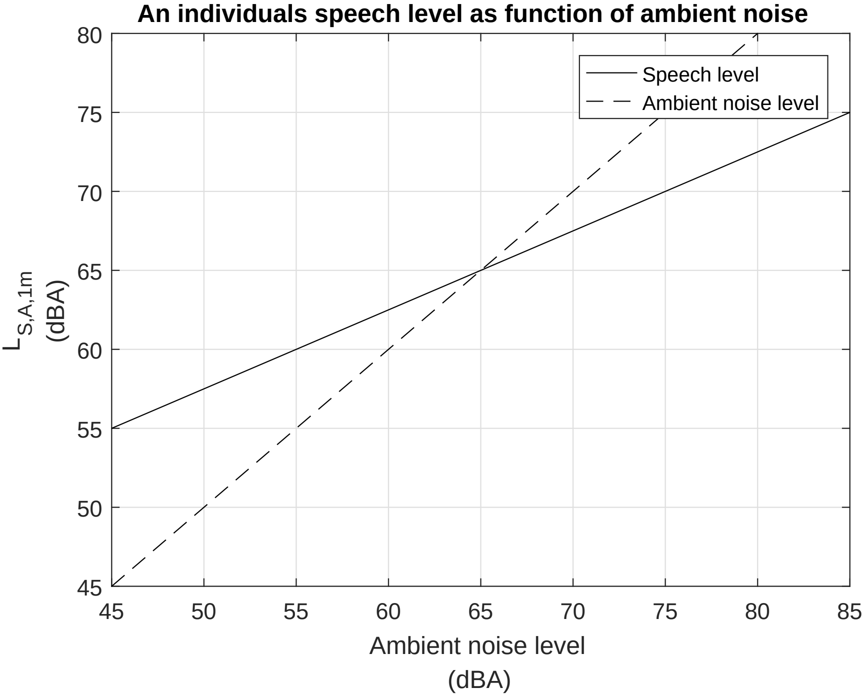


Figure 3.2: Speech level as a function of ambient noise

shows, when the ambient noise is 65 dBA an average person will be speaking at a level of 65 dBA as well. This means the SNR between noise and speech will be perceived as zero, at 1 meter from the speaker. When the noise is 71 dBA, the SNR is -3 dB which is defined as a minimum 3.4, for speech intelligibility. In this model, the only parameter to be decided is the slope of the increase, i.e. how big the increase in vocal effort is, per increase in noise. Rindel choses $c = 0.5$, but different studies have reported values in the range from 0.2 to 0.9.

Non-linear Lombard model

A more complex model of the Lombard effect has been proposed (Hodgson et al., 2007). Hodgson introduces more parameters in the model and with mathematical optimization techniques arrived at results that agreed with the available data. This non-linear model proposes that the increase in vocal effort varies at through the vocal range of the talker. That is, the response to noise is different according to how loud a person is speaking. The model includes several parameters (variable name in parenthesis):

- minimum vocal level ($L_{S,A,inquiet}$) (relaxed conversation in quiet surroundings)
- maximum vocal level
- the slope at the mid point between these two extremes
- how large the difference between the two extremes is ($asym$)
- a constant factor ($scale$)

The mathematical equation for the model is:

$$L_{S,A} = L_{S,A,inquiet} + \frac{asym}{[1 + exp\left[\left(xmid - L_{N,A}\right)/scale\right]]} \tag{3.7}$$

where $L_{S,A}$ is the resulting speech level and $L_{N,A}$ is the ambient noise level. An illustration of the general shape of the increase in speech level over increased noise is shown in figure 3.3. The overall shape of the curve shows that the increase in speech level is more gradual at the extremes of the range. People compensate less at the beginning of the Lombard effect (around $L_{N,A} = 45dBA$) and when they get close to their maximum speech level. The increase is greatest at the mid point of the vocal range. This midpoint is used to define the value of the Lombard slope.

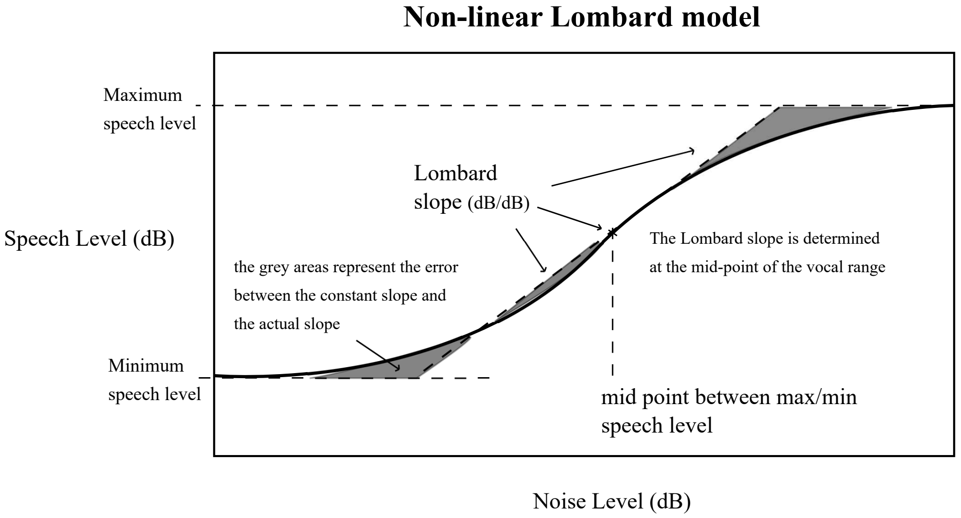


Figure 3.3: The nonlinear Lombard model

Lombard effect in ISO 9921

The ISO 9921 standard on Assessment of Speech Communication also includes the Lombard effect and indicates the variability between speakers. The relation between the range of vocal effort (equivalent continuous speech sound level) and the ambient-noise level at the speakers position is plotted in Figure 3.4 .

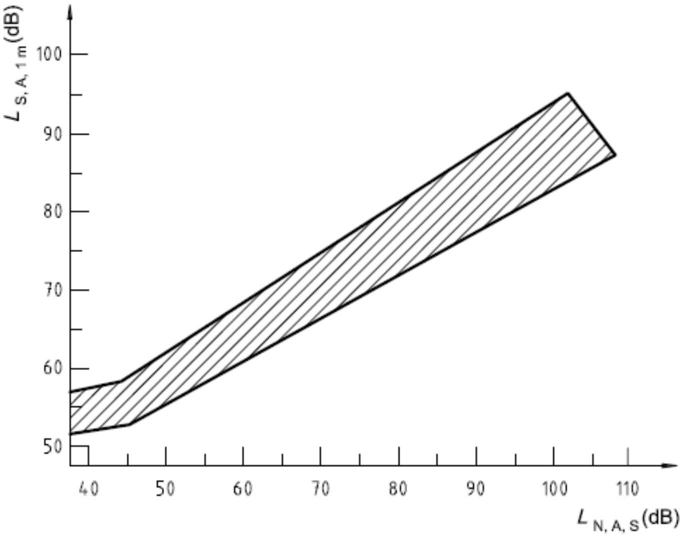


Figure 3.4: Relation between the vocal effort and ambient-noise level (Illustration from ISO 9921)

where the x-axis is A-weighted ambient noise level $L_{N,A,S}$ and the y-axis is A-weighted speech level at 1 meter distance from the speaker's mouth, $L_{S,A,1m}$
The hatched area indicates the variability of the Lombard effect among speakers.

3.4 Recommended speech and noise levels

There exists several methods for assessment of speech intelligibility. They are all based on the principle of comparing speech level to noise level, at the listener's ear (Lazarus, 1987). The accepted boundaries for background noise and required speech levels is presented in the following subsections.

3.4.1 recommended vocal effort

The level of the speech signal depends on the vocal effort of the speaker. The term *vocal effort* refers to how much effort it takes an average person to generate a speech level. This is tied in with comfort and the overall quality of the dining experience. If the ambient noise level induces speech levels above 60 dBA (”Normal”) the speakers will grow tired after a while.

The vocal effort is expressed by the equivalent continuous A-weighted sound-pressure level of speech measured at a distance of 1 m in front of the mouth, $L_{S,A,1m}$, $(dB\ re\ 20\ \mu Pa)$. The relation between vocal effort and the corresponding speech level at 1m is summarized in figure 3.5:

Vocal effort	**dB**	$L_{S,A,1m}$
Very loud	78	speech intelligibility is greatly reduced at speech levels > 75 dBA
Loud	72	
Raised	66	
Normal	60	**Recommended range** for prolonged person-to-person communication
Relaxed	54	

Figure 3.5: Vocal effort of a male speaker and related A-weighted speech level.

The standard states

> ”In situations of a relaxed type of communication, for example, occurring in offices, during meetings, lectures and performances, which take place over a longer period of time, a good level of intelligibility is recommended allowing for a normal vocal effort. ”

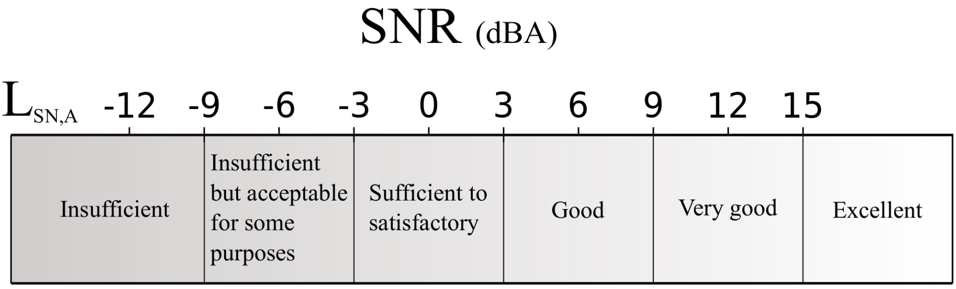
ISO 9921 defines Normal vocal effort for a male speaker to be $L_{S,A,1m} = 60\, dB\ (re\ 20\ \mu Pa)$. In other words, the ambient noise level should not be so loud that it induces speech above 60 dBA. Another point of interest; loud speech, above the level of $L_{S,A,1m} = 75\ dB$ is significantly more difficult to understand.

3.4.2 Speech-to-noise ratio (SNR)

The recommended SNR for verbal communication was summarized and presented by Lazarus (1987). A summary is given in figure 3.6.

3.5 Correcting for distance

Speech level is defines by the sound pressure level measured at 1 meter from the speakers mouth. The level that matters is the level at the *listener’s ear*. If the talker and listener are sitting 1 meter apart these levels are the same and no correction is needed. If however, the distance is something other than 1 meter, the estimated voice level will not represent the



$L_{SN,A}$ is the ratio between A-weighted speech level and A-weighted noise level

Figure 3.6: Recommended speech-to-noise ratio

level perceived by the listener. This is due to the way sound propagates and loses intensity over distance.

For distances other than the reference distance of $r_0 = 1\ m$, equation 3.8 can be used to determine $L_{S,A,L}$ (speech level at listeners ear:

$$L_{S,A,L} = L_{S,A,1m} - 20\ lg\frac{r}{r_0} \tag{3.8}$$

where *r* is the distance in meters between the speaker and listener and $r_0 = 1\ m$
This corresponds to a 6 dB drop in speech level at the listeners position, for each doubling of distance. This relation is valid indoor and outdoor, for distances up to 2 m. For conditions with a reverberation time smaller than 2 s at 500 Hz, a maximum distance of 8 m is valid (ISO, 2003). Some possible cross-table distances and their corrections (in dB) are shown in table 3.1.

Consider a restaurant scenario where the measured ambient noise level is 67 dBA and the average cross-table distance is 1.5 meter. The ambient noise will induce the average speaker to produce a speech level of $L_{S,A} = 66\ dBA$ (from eq. 3.6 on page 12). But the speech level at 1.5 meter would be 62.5 dBA (from eq. 3.8). The resulting ratio between speech and noise becomes $SNR = 62.5 - 66 = -3.5$. This is *insufficient* (ISO, 2003) for prolonged verbal communication. If the cross-table distance had been 1 meter (and

Cross-table
distance correction

r	Correction (dB)	
0.5	+6	
1	0	
1.5	-3.5	
2	-6	
2.5	-8	only valid for $T < 2s@500Hz$
3	-9.5	only valid for $T < 2s@500Hz$

Table 3.1: Various distances and their corrections

everything else the same) the result would have been $SNR = -1$ which is sufficient. Not a small difference in outcome!

However, in real-life situations there will be other factors involved and things won't be quite as clear-cut. For one, the Lombard effect is partly due to a learned response to the listener's need. I.e. the talker will instinctively raise their vocal effort to maintain a comfortable SNR for the listener (van Heusden et al., 1979). The interplay between these factors has implications when measurements at low occupancy are used to estimate levels at near-full occupancy. This will be discussed further in the analysis chapter.

3.6 Acoustic Capacity

The concept of *Acoustic Capacity* (AC) is a way to simplify complex room acoustic parameters down to a single number rating (Rindel, 2010). That number, the acoustic capacity of a venue, is the maximum number of people it can hold and still maintain a sufficient quality of verbal communication. As before, the measure of that acoustic quality is the SNR. The foundation for the concept comes from combining equation 3.2 with the accepted limits of verbal communication (3.6) and the results of Sabine. So, when room volume and reverberation time is known, the maximum number of guests can be calculated:

$$N_{max} \approxeq \frac{V}{20 \times T} \tag{3.9}$$

where V is the volume in m^3 and T is the reverberation time in seconds in furnished but unoccupied state at mid frequencies (500 Hz to 1000 Hz). There are some important assumptions made in the equation, related to the parameter *group size* (discussed in detail in the section below), and absorption per person. For equation 3.9 Rindel assumes that every guest in the room has an average absorption area of $A_p = 0.35\ m^2$ and the group size is $g = 3.5$.

3.6.1 Group size

The parameter called *groupsize* is defined as the average number of people per speaking person.

$$g = \frac{N}{N_s} \tag{3.10}$$

where N is the total number of people present in the room and N_s is the number of people speaking at the same time.

The group size can be thought of as a describing the excitement in the room, or the arousal (Rindel, 2010). It holds a lot the social information about the type of gathering, type of people present etc. A high group size (e.g. $g = 4$) signifies a calm and relaxed gathering. A low group size (e.g. $g = 2$) suggests a lively party. In most cases, as the evening progresses the group size will change. This is especially true if alcohol is served. Determining the correct group size is difficult and one of the main uncertainties of the concept (Rindel, 2010).

The concept of acoustic capacity involves 3 main assumptions:

- The number of people per speaking person (group size)
- The average distance between people (1 meter)
- The Lombard slope

So the calculated acoustic capacity is only accurate provided the acoustician guessed the correct table shapes, the type of social context it will be used and the true nature of the Lombard effect. These are big questions that need to be answered with more research. The data surrounding the Lombard effect is considered statistically weak, by several authors ((Rindel, 2010),(Lazarus, 1987), (Hodgson et al., 2007)).

In a way, the acoustic capacity is the opposite of what is proposed in this thesis. AC requires room acoustic parameters (V,T,A) as inputs and tries to determine the ambient noise level (produced by a given number of guests). The cutoff point for number of guests is the SNR. This is all information useful to the restaurant owner. In the simple method proposed here however, the ambient noise level is the *only* quantity measured. The output is a an estimation of the SNR, at 80% occupancy. And it's aimed at the potential guests, not the owners. So it's a different strategy working towards the same problem, assessing a restaurant's acoustic quality.

Some of the guess-work can be eliminated in an assessment situation, by observation. A quick overview of the restaurant and its guests can give some helpful additional information. For one, it's not realistic to expect the restaurant to be at maximum capacity when the measurement is done. It's more reasonable to assume 80% occupancy as the maximum, and estimate that noise level. Secondly, what is the group size? Are there observable clues that can make the determination of g more precise? By observing table shapes (long tables, square etc) the group size uncertainty can be adressed. And third, what is the distance between people? Is the assumption of 1 meter correct? In the analysis chapter, these questions will be discussed further.

3.7 Speech Intelligibility Index

The speech intelligibility index is an objective measure that is highly correlated with speech intelligibility (ANSI-S3.5, 1997). It's a more advanced calculation method than the SNR method or AC. SII requires a more detailed measurement of the noise spectrum and is a complex calculation method. In the framework of this thesis, the SII is an opportunity to view the acoustics from a different perspective. Instead of projecting what the noise level will be at 80% occupancy, it can show the speech level required . By calculating the SII from all 4 vocal efforts, one can predict the intelligibility for each one. So, it is an opportunity to estimate what vocal effort is needed to maintain verbal communication. It's included in the thesis as a way to solidify the accuracy of the simpler method. The exact calculation procedure is explained in the ANSI 3.5-1997 standard, but the user must make several choices based on the situation. In the following subsections, the relevant choices for restaurant acoustics are discussed.

3.7.1 Overview of the SII

Not all frequency ranges contribute equally to speech comprehension. For higher accuracy, the speech cues must be broken down into frequency ranges and two questions must be answered. First, how audible is the speech in a given frequency band? Secondly, how important is that band, for speech intelligibility? By first computing the speech-to-noise ratio in all bands and then weighing their relative importance, a score is produced. This score is a number between 0.0 and 1.0, called the *speech intelligibility index* (SII) . The speech intelligibility index is a measure highly correlated with speech intelligibility. The SII was defined in ANSI S3.5-1997 by the Acoustical Society of America. It indicates the proportion of total speech cues that are correctly understood, by the listener. The maximum SII value, 1.0 signifies that all speech cues are understood, the minimum SII score of 0.0 means no speech cues are delivered to the listener. Likewise, an SII of 0.5 means half the speech cues are available to the listener.

3.7.2 Adapting the SII to restaurant acoustics

The ANSI 3.5 standard provides a framework that covers a wide range of situations. It defines a *reference communication situation* (section 3.5) that can be adapted for each specific case. When the SII is used to evaluate restaurants, the user must make some choices and assumptions. Table 3.2 shows the necessary adaptations.

	SII	**Restaurant**	**ANSI 3.5**
1	Audio	Audio-visual	B.1 (Annex B)
2	monaural	binaural	5.1.5, 5.2.4
3	speech level is measured	speech level is assumed	5.1.3
4	noise measured at listeners pos.	noise meas. in diffuse field	3.15
5	No hearing loss	Hearing loss, age 30 and above	

Table 3.2: SII specific to Restaurant acoustics

1. Audio-visual The ANSI 3.5 is based on audio-only listening tests and the relative importance of frequency bands (expressed in the Importance function) were determined from these. However, visual cues from observing the talker's lips and face contribute to speech intelligibility. For situations where speaker and listener have visual contact, the SII does not directly apply. The standard states that either (1) new importance functions must be developed in each case or (2) an approximate "audio-visual SII" can be derived, from the computed audio-only SII. For a proposed method of evaluating restaurants, option (1) is far beyond the scope of any practical application. For one, it would require a large number of listening tests for each restaurant. Option (2) is to first calculate the audio-only SII and then translate it into an audio-visual SII (S_{av}), using the following equation:

$$S_{av} = b + cS \tag{3.11}$$

where S is the computed audio-only SII. For S not greater than 0.2, $b = 0.1$ and $c = 1.5$ For S greater than 0.2, $b = 0.25$ and $c = 0.75$.

Equation 3.2 shows that audio-visual cues are more significant in situations where speech intelligibility is low. A computed $S = 0.3$ (poor) will translate to an audio-visual $S_{av} = 0.475$ (adequate), an improvement of almost 50%. Whereas $S = 0.8$ (good) becomes $S_{av} = 8.5$ ($\approxeq 6\%$ improvement).
The necessary condition for equation 3.2 is ”good visibility of visual cues” . Visibility depends on various factors such as lighting conditions and table shapes etc., and will vary between restaurants. However, it can be argued that most, if not all restaurants will have good visibility, due to the social context. Put simply, when people sit down for a meal, they expect to be able to see each other. So, the suggested method will assume good visibility and factor in audio-visual cues, using equation 3.2.

2. Binaural The reference communication situation in ANSI 3.5 is monaural. In a restaurant, it should be expected that the majority of people have normal hearing in both ears (binaural). In such cases, the equivalent hearing threshold must be adjusted by -1.7 dB in each frequency band (see 5.1.5, 5.2.4).

3. Speech level One of the inputs for the SII computation is the speech spectrum level of the talkers. Clause 5.1-5.3 offers methodology for direct measurement/estimation of noise and speech levels. However, measurements of a stable and accurate speech spectrum level requires a large group of talkers (at least 20 talkers and 30-second speech samples from each are recommended). Therefore, the standard recommends using a statistical dataset, the *standard speech spectrum level* (E_i).
Table 3 in ANSI S3.5 presents the ”standard speech spectrum level for stated vocal effort” , for the 1/3 octave band procedure. The spectra represent an average of adults (male and female), for ”average speech” . The spectra are shown in Figure 3.7 below. An interesting phenomenon can be observed. When people raise their voice, an increase in loudness is not the only change. The spectral content changes and the concentration of energy shifts slightly upwards in frequency. However, for a vocal effort ’normal to raised’ the speech will have a similar spectrum (Lazarus, 1986).
The SII calculation is based on free-field levels, measured at the center of the listener’s head (when the listener is temporary absent). The standard speech spectra are measured 1m from the talker’s lips. If the distance between guests is 1m, the equivalent speech spectrum level ($E_i^{'}$) is the same as the standard speech spectrum level E_i. If not, the equivalent speech spectrum level can be derived from the following equation:

$$E_i^{'} = E_i - 20log\frac{d}{d_0} \tag{3.12}$$

where d is the distance between guests and d_0 is the reference distance of 1m.

4. Noise level In the reference communication situation, the noise spectrum level is measured at the center of the listener’s forehead. In a restaurant, it is not practical to calculate an SII for each listening position. A diffuse field measurement of the noise is the realistic option. This measurement is then converted to the level at the listener’s eardrum, using the free-field-to-eardrum transfer function (See 3.29).

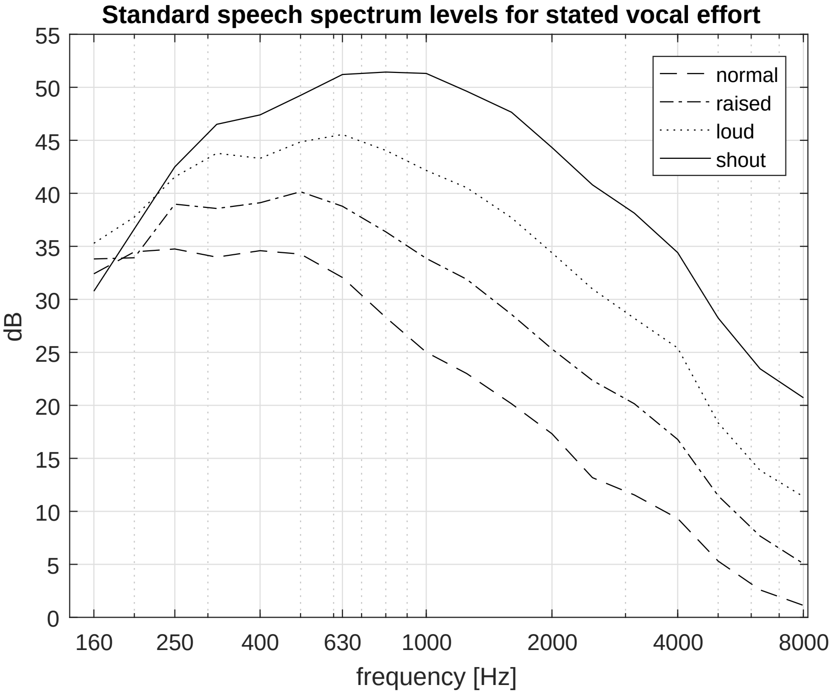


Figure 3.7: Standard speech spectrum levels

5 For the SII calculation, unless audiograms are available the equivalent hearing loss is set to 0 in all bands. This is statistically correct for people aged 18-30. Since many restaurants have older guests, it will be more accurate to calculate a separate SII for various age groups. This will take into account age-related hearing loss.

3.7.3 Calculating the SII

The Acoustical Society of America offers program for SII calculations. Hannes Msch and Pat Zurek, both members of ASA Working Group S3-79, have developed a Matlab script for calculating the SII. It is available at http://www.sii.to/html/programs.html.
The SII calculation is based on free-field levels, measured at the center of the listener's head (when the listener is temporary absent). The measured 1/3-octave band levels must be converted into spectrum levels. Table 3 of ANSI 3.5 proves the correct bandwidth corrections. The measured values must also be converted from free-field to eardrum levels. Table 3 of ANSI 3.5 provides the correct transfer function.

The Acoustical Society of America offers program for SII calculations. Hannes Msch and Pat Zurek, both members of ASA Working Group S3-79, have developed a Matlab script for calculating the SII. It is available at http://www.sii.to/html/programs.html.
ANSI 3.5 defines an SII of 0.75 or above as ”good” . An SII of 0.45 or below is defined as ”poor” . Since the reference communication situation is monaural and audio-only, the SII

must be converted to a binaural audio-visual SII.

3.7.4 Transferring SII to speech intelligibility

In ANSI S3.5-1997, the SII is defined to be ”highly correlated with speech intelligibility under adverse conditions” . However, it also states that the calculated SII is not directly converted to any corresponding speech intelligibility score. There are other factors involved which influence the intelligibility of the received speech. It is therefore necessary to develop a transfer function that takes into account specific situations. The content of the message (linguistic, semantic and syntactic) and the proficiency of the listener will influence this step. The relative importance of frequency bands are different for various types of speech. For greater accuracy, the SII should be calculated using a band importance function characteristic of the actual speech material in a situation. However, this is not necessary in conditions where the equivalent noise spectrum level and the equivalent speech spectrum level are roughly parallel. In the context of restaurant acoustics, this is indeed the case.

3.7.5 Uncertainty of the SII for restaurant assessment

The SII is based on comparing speech level to noise level, factoring in hearing ability and the content of the speech. This covers a lot of important factors, but there is a drawback. Unless one is prepared to do the detailed testing required, the only option is to rely on statistical models, for all but the ambient noise. Average speech level, normal hearing and average speech content (syntax and semantic content). The only factor measured on site is the ambient noise. So is this method necessarily more precise than the simple SNR of A-weighted noise and speech levels? That would depend on the accuracy of the statistical models given in the standard. It is definitely a possible weakness, and questions have been raised about the speech model in particular [citation kommer]
Also, the SII is accurate for static noise, but less so when the noise fluctuates (Rhebergen and Versfeld, 2005). Fluctuating ambient noise is usually the case in restaurants, with the ebb and flow of conversations combined with random mechanical noises (chairs and cutlery etc).

Another uncertainty is regarding the importance function. There are two available importance functions that fit the situation. ”Standard” or ”Short passages” (also called ”Continuous speech”) and it’s difficult to interpret from the standard which one would give the most accurate prediction of speech intelligibility, in a restaurant.
Overall, the SII is used in the method proposed in this thesis as a ”second opinion” for the simpler ”SNR of dBA” test. A way to validate and check the simpler method. But since the noise spectrum consists largely of voices the two methods should most likely yield similar results. If, on the other hand, the ambient noise had been very different from speech, say machine noises at a factory, the SII would be better suited to assess the situation. It would be able to determine if the important frequency bands for consonants were being masked by the noise.

3.8 Privacy

In restaurants, the ambient noise has a masking-function that gives the guests a sense of *privacy*. Knowing that people nearby can understand your conversation makes people uneasy. In order to maintain privacy, it's important that the SNR between the speech from one's own table and the ambient noise at *surrounding* tables is below -9 dB (Long, 2005). This requirement puts an upper limit on how much absorption per table and a minimum distance between tables. In practice, the problem is usually not too much absorption, but too little. And since this evaluation method is aimed at providing information about the restaurant *as is*, this topic will not be discussed further.

If however a restaurant decides to improve their acoustic situation, the acoustician should be aware that there are limits to how much absorption can be installed. At some point, the only improvement is to lower the number of guests.

Chapter 4

Measurements

Measurements have been done in a single restaurant, to compare the different levels of analysis. In this chapter the results are presented. The measurements were done according to standard engineering practices.

4.1 Test case: ”Restaurant A”

The ambient noise in Restaurant A was measured when the restaurant was at about 60% capacity. The 1/3-octave band levels were measured in 5 positions (diffuse-field). Each position was logged for 2 minutes with the RTA module in Audio Tools. The equivalent continous noise level is calculated and stored each second, based on L_n recorded each 0.1 second. The average value in each band, for the whole time series, is calculated directly by the software. The results are displayed in figure 4.1 An average of these 5 positions is shown in figure 4.2.

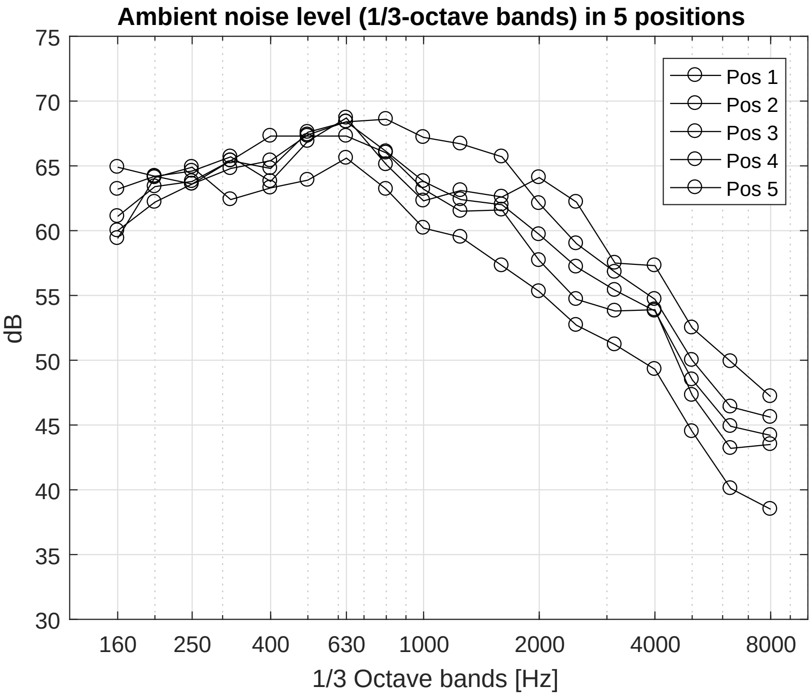


Figure 4.1: Ambient noise level in Restaurant A

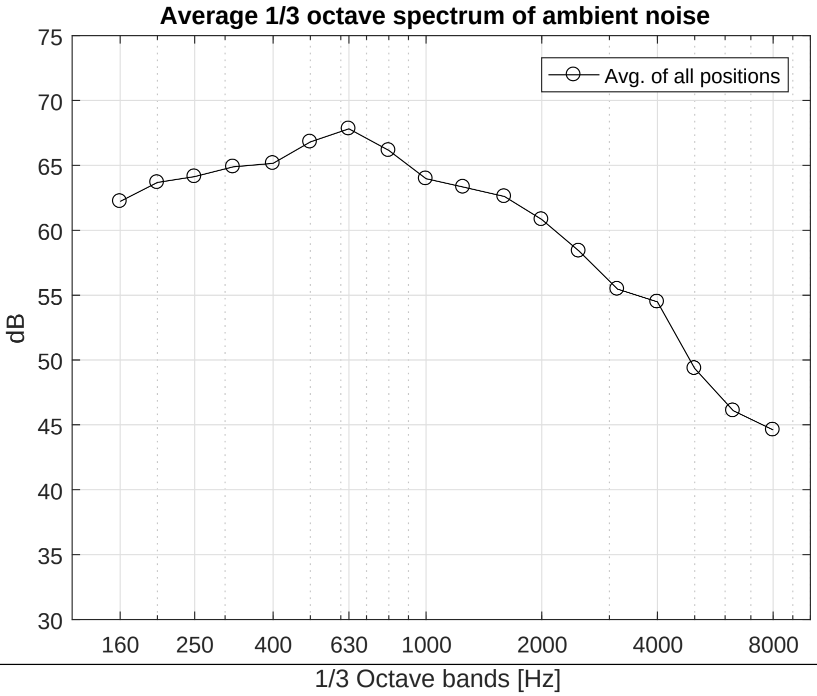


Figure 4.2: Averaged ambient noise spectrum in Restaurant A

The application also calculates the equivalent A-weighted level from the time series. The A-weighted equivalent noise levels, $L_{N,A}$, in each of the five positions are shown in table 4.1 The large variation in level is possibly due to the layout of the restaurant. It has a large floor area and a sound system at one side of the room. Sound system speakers are at the entrance to the restaurant, slightly away from the tables. Position 2, the least noisy position was at the opposite end of the restaurant from the sound system, with empty tables nearby.

Position	$L_{N,A}$ (dBA)
1	75.6
2	70.6
3	73.2
4	73.7
5	74.4

Table 4.1: The A-weighted equivalent ambient noise levels of restaurant A

The arithmetic average of the values is 73.5 dB.

For the SII calculation, the 1/3-octave band values must be converted into *spectrum levels*, i.e. the energy in a given 1 Hz band that corresponds to the measured value. The details for this can be found in ANSI 3.5, Table 3.

In figure 4.3 the average noise spectrum is plotted against the four statistical speech spectra from ANSI 3.5. ANSI 3.5 uses the symbol $N^{'}$ for the equivalent spectrum level of noise at a given position, that would have been measured in the reference communication situation.

4.1.1 Room acoustic parameters

Room volume: V = 765 m^3

Reverberation time: T = 0.6 s in furnished but unoccupied state at mid frequencies (500 Hz to 1000 Hz). The reverberation time was measured in accordance with ISO 3382-2:2008, using balloon popping noises as the impulse, in unoccupied state. The noise level unoccupied was 40.6 dBA, mostly due to the ventilation system.

4.2 Equipment

4.2.1 Audio Tools

The measurements were done with an iOS smart device running Audio Tools, made by Studio Six Digital. Audio Tools is a platform that provides most of the measurements found in typical audio analyzers. The equivalent continuous noise level were measured using the RTA module in Audio Tools, which meets all ANSI and ISO Class 1 standards for filter-based RTA measurements. This module calculates L_n at 0.1s resolution, which matches many traditional logging sound level meters.

The application is available at **http://studiosixdigital.com/audiotools/**

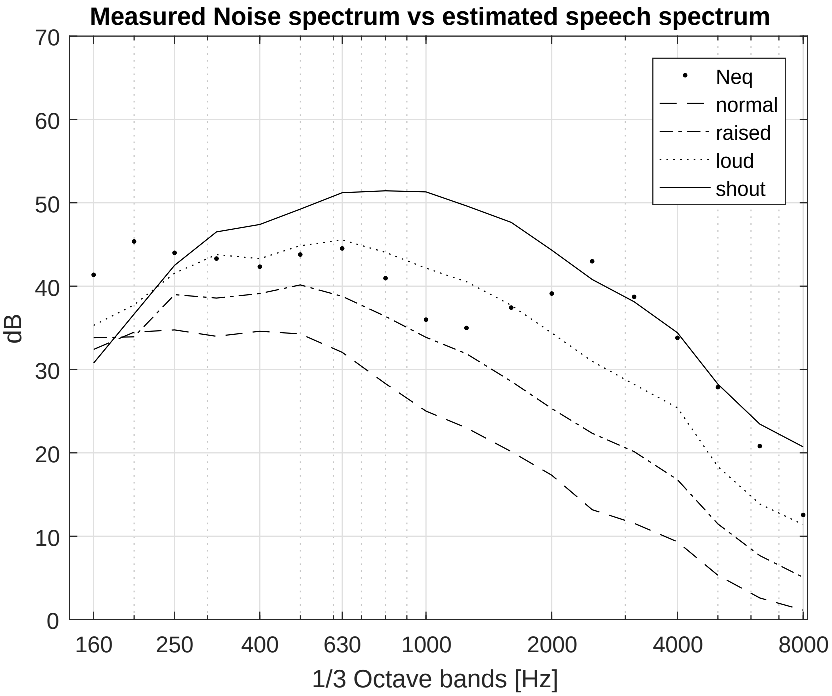


Figure 4.3: Measured noise spectrum in Restaurant A

4.2.2 iMic 436

The i436 is a calibrated measurement microphone that complies with the IEC 61672 Class 2 sound level meter standard.

Chapter 5

Analysis

5.1 The problem of averages

Many eating establishments have acoustic conditions that vary greatly. There might be a larger, ”main” area of the EE surrounded by different sections. Areas with booths, for instance or even different rooms (coupled or uncoupled). There might be a bar area with tables nearby (where music might be playing). Each of these areas will have different acoustical properties and noise levels. This begs the question - is it appropriate to describe that EE’s acoustic quality by an average of all the areas?

Let’s imagine a scenario where an EE has two distinct sections, one with ”good” acoustics and with ”bad” . If all guests present are polled at once, the average satisfaction of the restaurant might come out as ”okay” . For an assessment aimed at informing the public about what to expect at that EE, this is a problem. Each individual guest has a 50% chance of having a bad experience and a 50% chance of being pleasantly surprised . Is the comfortable noise level in the next room a plus for the guests in the ”bad” section? For an individual guest, the answer is obviously negative. As an overall indicator of customer satisfaction, the ”okay” rating might be accurate, especially if the ”good” section is usually a little more crowded than the ”bad” section.

Since this assessment method aims at providing a tool to be used a large scale, to effect change in society, the score it produces must make sense to the customers, the potential guests of that EE. It’s less important to estimate the average satisfaction level and more important to give information about what experience a person will have, at a given EE. So, it then makes sense to include the various differing spaces, in the test. At the same time, if this is done in a very detailed way it will complicate the procedure. So a balance between accuracy and simplicity must be found.

Suggested here is a simple statistical approach: If people are randomly assigned to various acoustically coupled sections, the average noise levels should be used. If the varying sections are un-coupled, roughly equal in size and/or have special uses, they should be evaluated separately and commented on. Figure 5.1 illustrates an example of a restaurant with various areas and spaces:

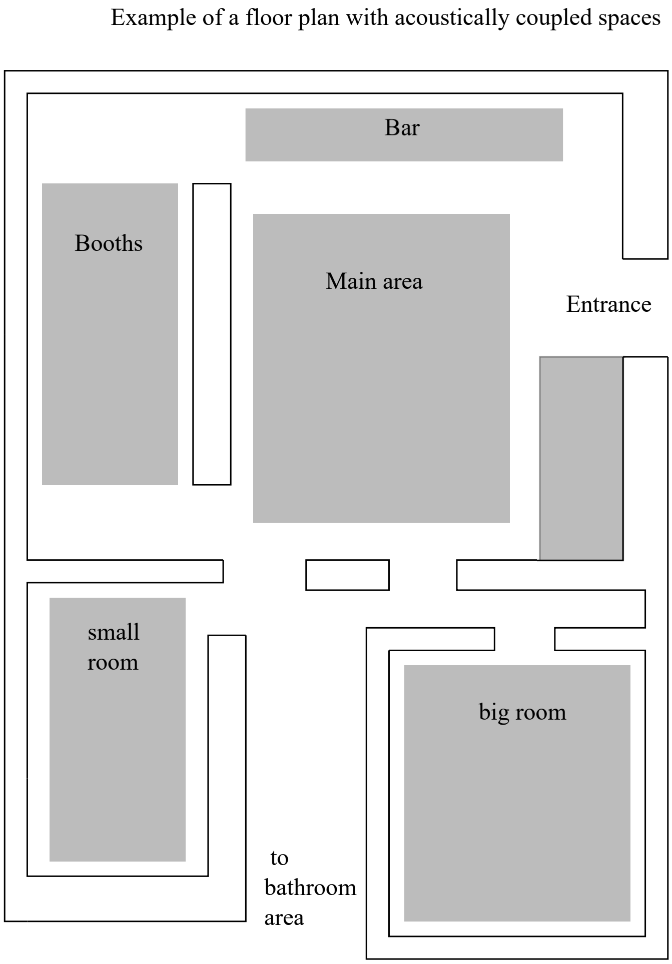


Figure 5.1: Illustration of a restaurant with acoustically coupled spaces

This EE has five areas with potentially different acoustic qualities, all connected via openings. In this situation it is complicated to predict how the different areas will influence each other, i.e how much noise will leak between them. So, estimating the noise level when all sections are full, based on a measurement at less than maximum, becomes less certain. This is a weakness of the proposed single-measurement method, because of the underlying diffuse-field assumption. However, some simplifications can be done. The main area, bar and booth section can be regarded as one. This would mean there are two rooms coupled with each other and the main area. So how could the acoustician proceed?

One option would be to assess each area separately. However, dividing the assessment into more than two sections could perhaps complicate matters, for the potential guests. For example, comments like ”the big room should be avoided if the restaurant is more than half-full, but the small room ....” . It might not be obvious to the average person what room or area the acoustician is referring to.

So in cases such as Figure 5.1, the best compromise might be to base the assessment on ambient noise levels in the ”main area” . The necessary assumption would then be that the rooms will fill up at an equal rate, i.e new guests would be distributed equally across the various sections. One would also assume that the noise level rises equally in each section, for each new arriving guest. Neither of these assumptions are likely to be true, but the inaccuracy they introduce is likely to be small. This is due to the partial separation between the different areas. The guests in the smaller areas contribute less (individually) to the overall noise level in the main area, than guests in the main area itself. So if for instance the acoustics of that ”big room” leads to a quicker rise in noise level *in there*, as the number of guests increase *in there*, the effect of that *in the main area* will be less drastic. So a reasonable approximation will most likely be to just measure the larger, main area and pretend each new guest will behave according to the known Lombard theory and contribute equally to the noise level. The crucial factor here is that the ”main area” is significantly larger than the other areas.

The main exception to this approach would be when an EE has two areas that are acoustically uncoupled, to a large degree. Examples would be a restaurant with a room upstairs, possibly used for special bookings (closed parties, larger groups etc). In that case, the separate area might deserve its own comment in the assessment. Especially if that section is promoted by the EE for such uses. In this scenario it would not make sense to use an average the two areas to represent the acoustical quality of the EE. Instead, separate measurement and assessment of the other space, with it’s own comment, is necessary.

5.2 Occupancy

Restaurants are rarely ever completely full of guests. A more realistic evaluation might be to set 80% occupancy as the maximum. So for instance, if the restaurant is only at 50% capacity during the measurement, that noise level should be used to estimate what the noise level will be at 80% occupancy. Roughly speaking, a doubling af the number of guests results in a 6 dB increase in $L_{N,A}$ (Rindel, 2010). Thus, if the restaurant is half full when $L_{N,A}$ is measured the predicted maximum level is 6 dB higher.

5.3 Table shapes and sizes

When acousticians try to predict what the noise level will be in a restaurant, one of the unknown factors is table shapes. Table shape matters in two ways: (1) It determines the physical distance between talker and listener. (2) It influences behavior in terms of how many conversations can happen simultaneously, in a group. Social norms vary between demographic groups and social settings, so precise prediction can be difficult. But since the restaurant can be observed *in use*, the distance between people can be known. This allows for some of the unknowns to be checked and corrections to the model can be made. One such factor that can be observed directly is *table shape*. The table shapes will affect the parameter *group size*. The number of conversations that can happen simultaneously will be different, for different table shapes. When four people sit together at a square table, it is reasonable to assume that they will be part of the same conversation. One person will be speaking at a time, with 3 people listening. This leads to the groupsize of $g = \frac{N}{N_s} = 4$, illustrated in figure 5.2

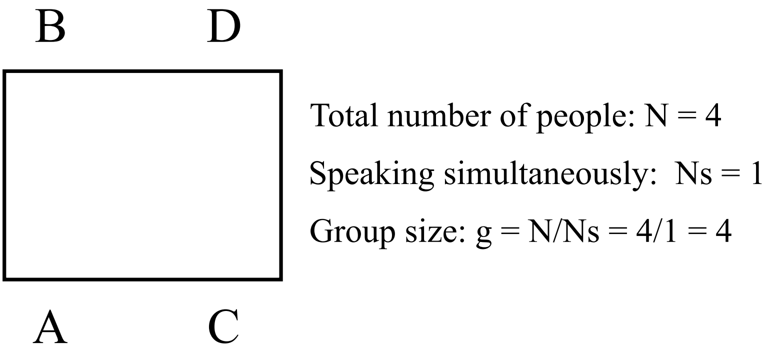


Figure 5.2: Square table

Now let's say there are 4 tables, each with 4 people, for a total number of $N = 16$. With this table layout, there will be four simultaneous speakers, $N_s = 4$.

The same is not true for a long table, with the same number of people present. If 16 people sit at a long rectangular table, as shown in figure 5.3, it's likely that more than 4 people will be speaking at once. Each individual has several options for conversation, and they will also be more distracted by other nearby conversations and thus speak louder.

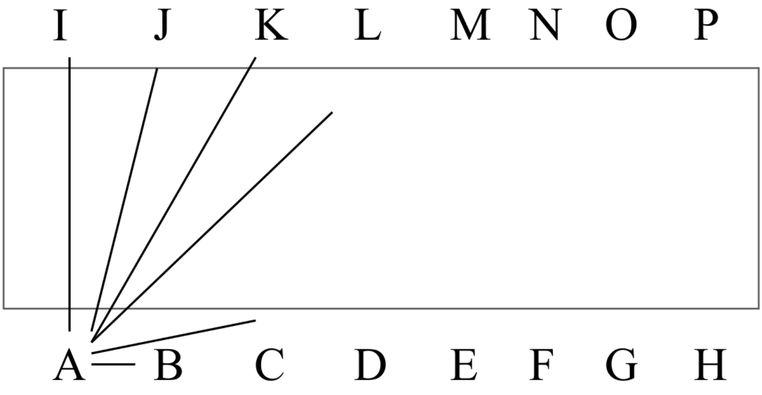


Figure 5.3: Long table

If there are several long tables, the group size will likely be lower than average. I.e. as number of guests increase, the noise level will escalate quicker than the model suggests. This happens because there are more people speaking simultaneously, meaning there are fewer people per speaking person (lower group size). So, the shapes of a restaurants tables matters. This has implications when an EEs acoustics are being assessed. If there are many long tables it is likely that the noise level will be higher than if the same number of seats were distributed across smaller tables.

Furthermore, in a typical situation not all tables will be full. The examples above are just to illustrate the difference. Realistically, some tables might have only 3 or 2 people. Also, sometimes people speak over each other, especially if alcohol is served. So a more sensible estimate has been suggested: $g = 3.5$ for a quiet bistro and as low as 2.5 for a more lively place (Rindel, 2010). Again, this behavior will vary between different types of establishments. Group size is the most difficult parameter to determine (Rindel, 2010) and probably the main uncertainty in the concept of acoustic capacity. So, the opportunity to minimize this uncertainty through observation should not be missed.

5.3.1 Distance between talker and listener

All the various methods for determining speech intelligibility are based on the same principle; comparing the speech level and the noise level, at the listeners ear. So the physical distance between people becomes highly influential on the calculated result. Again, for prediction purposes an assumption has to be made. And again, for assessment, this can actually be observed directly. How big are the tables? Specifically, what is the distance between speaker, across table? Equation 3.8 shows how the SNR can be recalculated for different distances. If the average cross-table distance is 1.5 m, the speech level at the listeners ear is:

$$L_{S,A,1.5m} = L_{S,A,1m} - 20\, lg \frac{1.5}{r_0} = L_{S,A,1m} - 3.5$$

This was discussed in section 3.5 on page 15, and some typical correction values were presented.
If a talker is speaking at a normal conversational level of $L_{S,A,1m} = 60dBA$, against an ambient noise level of 60 dB, the SNR drops to -3.5 dB, which is insufficient for prolonged verbal communication. Furthermore, a longer across table distance than 1 meter means the noise will escalate more quickly, for each additional guest. This will happen because each individual is speaking at a louder speech level than the model assumes, so fewer individuals are needed to set the effect in motion.

For the proposed single measurement method, this has an important implication; At a restaurant with longer cross table distances, a given ambient noise level can be reached by fewer people. I.e the increasing spiral of noise will kick in at a lower occupancy rate in a restaurant with 1.5 m cross table distance than at one with 1 m. So, when a measurement is done at a lower occupancy rate than 80%, the estimated maximum noise level will be higher if the distance is not the standard of 1 m.

How to factor this into the assessment is a difficult question. It is still not known how much of the increase in vocal intensity is due to an automatic reflex (Lombard effect) or due to a learned response to listener's needs (Jr. et al., 1989). But it is known that *some* of it will be based on prior experience with speaking in noise. The speaker will instinctively raise their vocal effort to compensate for the distance, which would help the SNR. But then the vocal effort needed will take more effort as well. In the example above, the vocal effort will be above the recommended range (ISO, 2003), deeming the restaurants acoustics to be of insufficient quality.

The suggested solution in this method is to use a slightly scaled down version of the corrections shown in 3.1. By lowering the correction values, ”some” of the learned response is factored in. I.e. the talker is assumed to pick up on visual cues and thereby raise their voice ”some”, to compensate for the longer distance. Knowing exactly how much ”some” is would require knowledge of how much of the Lombard effect is involuntary and how much is due to experience. This is not known at present time, so some guesswork is needed. Suggested here is a ratio of 50/50 between the involuntary response and the ”natural compensation” due to experience. And so the corrections from table 3.1 becomes table 5.1. This scaling down from the theoretical value factors in some of the known complexity of the Lombard effect. Whether or not choice of scaling it down by half is the most appropriate is not possible to know without statistical data. But the method will, by logic, be more accurate than if the ”natural compensation” is not factored in at all.

Suggested cross table distance correction

r	Correction (dB)	
0.5	+3	
1	0	
1.5	-1.75	
2	-3	
2.5	-4	only valid for $T < 2s@500Hz$
3	-4.75	only valid for $T < 2s@500Hz$

Table 5.1: Suggested distance corrections

The ambient noise in an EE is not only a negative. It also has the important and positive function of *masking*. In a multitalker situation, a common concern for people is that their own conversation must feel private. The reverberation time and distance between tables are the two main factors here. It is possible to have too much absorption in a restaurant, which leads to an SNR that is too high at surrounding tables. I.e the speech from other tables isn’t sufficiently masked by the ambient noise, at ones own table (and vice versa). Then, the content of speech from surrounding tables can be comprehended. This is not only annoying and distracting, it also has a psychological effect on the guests. They feel a lack of *privacy*, as their own speech would probably be understood at the other tables.

Encountering a situation where the problem is too *much* absorption might not happen often. A more likely scenario is that there are simply too many guests compared to the room’s volume (V). Then, the acoustician might try to lower the ambient noise by increasing the total absorption area (A) to compensate. If this approach is taken to far, it will lead to a lacking sense of privacy. The reverberation time can become too short for sufficient masking of speech, between tables. This will be especially apparent at mid/low occupancy rates. But the problem here is really about the number of guests and the room volume.

5.4 Classification

Defining the limits between the classes is a crucial step. The ISO9921 standard is used here as a basis for classification. In the standard, an SNR between -3 and +3 is defined as ”sufficient to satisfactory” . Since this evaluation is meant for communicating to anybody, this class has been renamed here to ”okay” . A corresponding neutral smiley icon is chosen, to communicate nonverbally. The ambient noise level in this class is in the range 59-71 dBA.

If the minimum SNR of -3 is to be achieved, the upper limit for ambient noise is 71 dBA (Rindel, 2010). From 71 dBA and up, the acoustic quality is defined as ”insufficient”, renamed here to ”bad” and a corresponding sad smiley.

Ambient noise levels below 59 dBA guarantee an SNR of +3 or better, are classified as ”good” . In this ambient noise, talkers can speak with normal vocal effort (ISO 2003) and still be understood. This is important for prolonged conversations. Figure 5.4 illustrates the smiley icons and the limits. The SNR is a comparison of the A-weighted speech level at 1m distance and the A-weighted ambient noise level in the restaurant.

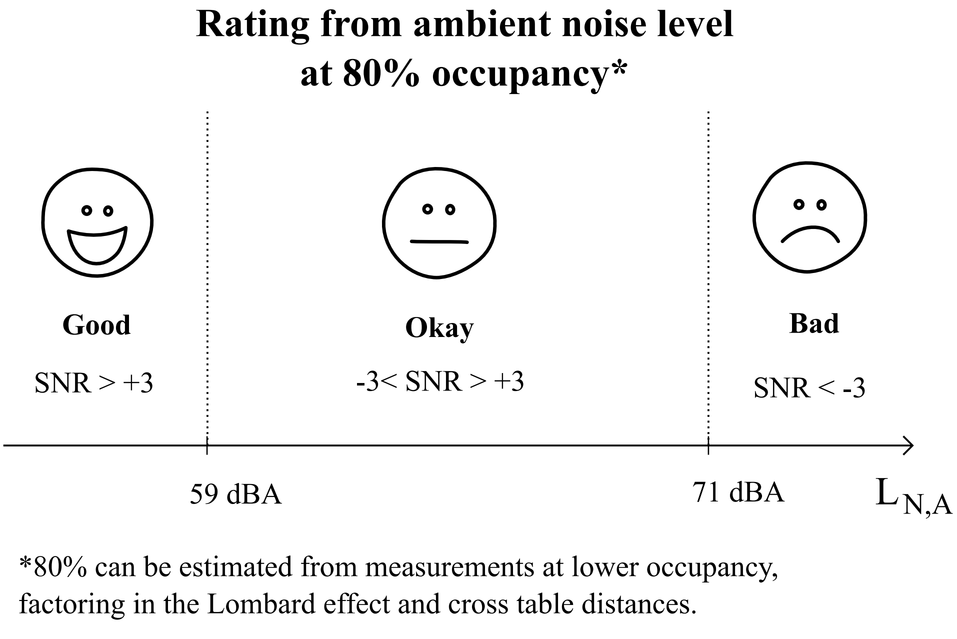


Figure 5.4: Ratings from ambient noise level at 80% occupancy

The limits have been chosen with the aim of producing an honest evaluation, from the guest’s perspective. So there is a possibility that very few restaurants will achieve the ”good” rating. But then again, that is a big part of the motivation for this method. Perhaps it could be argued that the classes should be divided in such away that most restaurants would fall in the middle category. But as long as the method provides an accurate representation of the situation, then so be it. The overall goal is to accurately predict the quality of the experience, from an acoustical viewpoint. Then people can decide for themselves how much it matters.

5.5 The single-measurement method

The single-measurement method suggested here requires three inputs:

1. Ambient noise level in occupied state (dBA and 1/3-octave spectrum)
2. Number of guests present during measurement and maximum available seats
3. Table shapes and placement

The acoustician must first make a judgment about the overall shape and layout of the room(s). Restaurant are rarely simple geometric shapes. Sometimes a restaurant has several spaces that may or may not be coupled acoustically. Some simplifications and assumptions are needed. The more complex the shape and layout is, the more the noise level and frequency spectrum will vary, spatially. And the less informative an average of all areas will be, to the potential guests.

After deciding how to evaluate the various areas, the ambient noise level can be measured in the appropriate places. Then, if the restaurant was not close to being full, knowledge of the Lombard effect can be used to estimate the ambient noise level at maximum occupancy (or 80% occupancy which might be more reasonable as restaurants are rarely completely full).

The method uses the known Lombard effect to enable measurements at lower occupancy be used to predict noise levels at high occupancy. The effect begins at ambient noise levels of around 45 dBA. As can be seen in 3.2 on page 12, the ambient noise eventually ”catches up” with the speakers (at around 65 dBA) for an SNR = 0. When the ambient noise has reached 71 dBA, the speech level will be around 68 dBA, for an SNR = -3 . This is the cutoff point for prolonged verbal communication (ISO, 2003). For ambient noise levels above 71 dBA, the quality of verbal communication becomes *insufficient*. However, at lower occupancy rates any restaurant can ”pass” this bar and have acceptable noise levels. By observing the occupancy rate during measurement the maximum level can be estimated with reasonable accuracy.

The distance across tables should also be factored in. The statistical speech levels used are defined as the level 1 meter in front of the speaker’s mouth. If the distance is larger it can have two consequences. One probable outcome is that talkers instinctively raise their voices, based on experience and visual cues from the listener. This means the escalation of noise starts earlier in the occupancy rate, the larger the tables are. It also means a higher vocal effort is needed, resulting in a less pleasant experience. The other possibility is that the person speaking does not raise their vocal effort, leading to a lower SNR at the listener’s ear. This hinders speech comprehension and requires more focus and concentration from the listener.

5.6 Overview of the method

Method overview

	Measurements on site	**Comment**	**Keywords**
1	determine if the acoustics can be evaluated as one single system	base assessment on 1 "main area" or divide	Diffuse field, coupled spaces
2	measure the ambient noise level, occupied	both in dBA and 1/3-octave spectrum	SNR, SII
3	observe across table distances	note deviations from 1m	Sound propagation
4	observe occupancy rate at time of measurement	# of occupied and unoccupied seats	
#	**Calculations**	**Comment**	**Keywords**
5	Estimate SNR for 80% occupancy		Cocktail party effect (Lombard effect)
7	correct for cross-table distances (if needed)	estimates the speech level perceived at the listeners ear	Sound propagation section 3.5, p. 15 table 5.1, p. 34
9	Set rating	recommended SNR are standardized in ISO 9921	SNR, ISO 9921

Table 5.2: Overview of the method

The measurement of the ambient noise level should be done in several locations around the restaurant and the results averaged. The assumption of a diffuse sound field is not always true and variations in noise level are a possibility. The A-weighted ambient noise level along with assumptions about speech levels (factoring in Lombard) will give a SNR for that particular restaurant (after necessary corrections in step 3,4 and 5).

A measurement of the frequency spectrum of the ambient noise will allow for a calculation of the Speech Intelligibility Index. There are assumptions of speech levels in the SII that may not be accurate, so the SII is not necessarily more valid than the A-weighted SNR obtained in step 1 (and 3,4,5).

5.7 Test case: "Restaurant A"

A restaurant has been measured with regards to reverberation time (unoccupied) and room volume. These measurements were used for the more complex and time-consuming methods of assessment. This was done in order to test the validity of the proposed single measurement method.

Restaurant A has a room volume of V = 765 m^3 and a reverberation time of T = 0.6 s. Inserting into equation 3.9 (page 17) gives $N_{max} = 63$. I.e. the maximum number of guests this restaurant can hold while maintaining an acceptable SNR is 63. Since the restaurant actually has almost 90 seats available for patrons, this will obviously present a problem. It also presents a problem in this real-world, simple method of assessment. What if the measurement was done on a day with only 63 guests. Or only 50? In a worst-case scenario, the acoustician makes the measurement on a day with only 63 guests and concludes that the acoustic quality is sufficient. So a good rating is given to a restaurant that does not deserve it. This is an important issue to address, if a single measurement is to be helpful. The ambient noise level in a half-full restaurant must be ”projected” up to what it would be, at 80% occupancy.

It becomes important at this stage of the evaluation to have a look around. There are two factors that complicate this ”projection” . ”Restaurant A” has several ”long tables”, which influences social interactions. This is a challenge in the more complex methods, as well. In the Acoustic Capacity calculation, it makes determining the *group size* parameter difficult. The reason is simple. This form of table allows for several conversations to happen at once, within a social group. It is reasonable to assume that the background noise will be even louder and escalate more quickly in a table arrangement of this type.

From the measured noise spectrum in Restaurant A, the SII was calculated according to ANSI 3.5-1997. The exact calculation procedure is explained in the standard. The necessary choices relevant for restaurant acoustics were discussed in Section 3.7 on page 18. The resulting SII score for the levels of vocal effort is shown in figure 5.5.

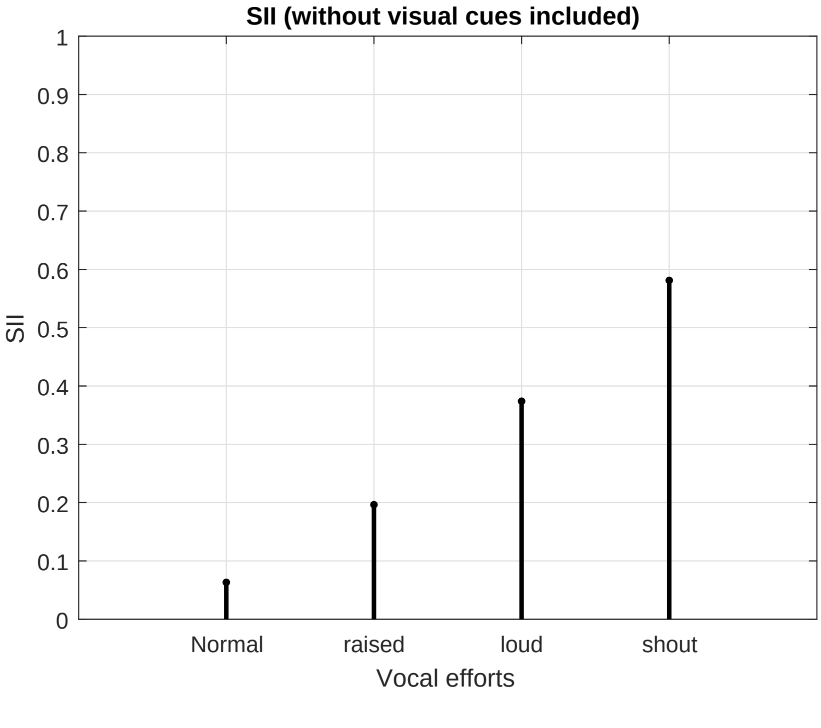


Figure 5.5: SII for Restaurant A, without corrections for visual cues

Visual cues play a large part in verbal communication. It’s reasonable to expect any EE to have sufficient lighting for visual cues to be included. As detailed in the chapter on theory, ANSI 3.5 includes an equation for transforming the audio-only SII to an audio-visual SII, S_{av}. The scores are displayed in figure 5.6.

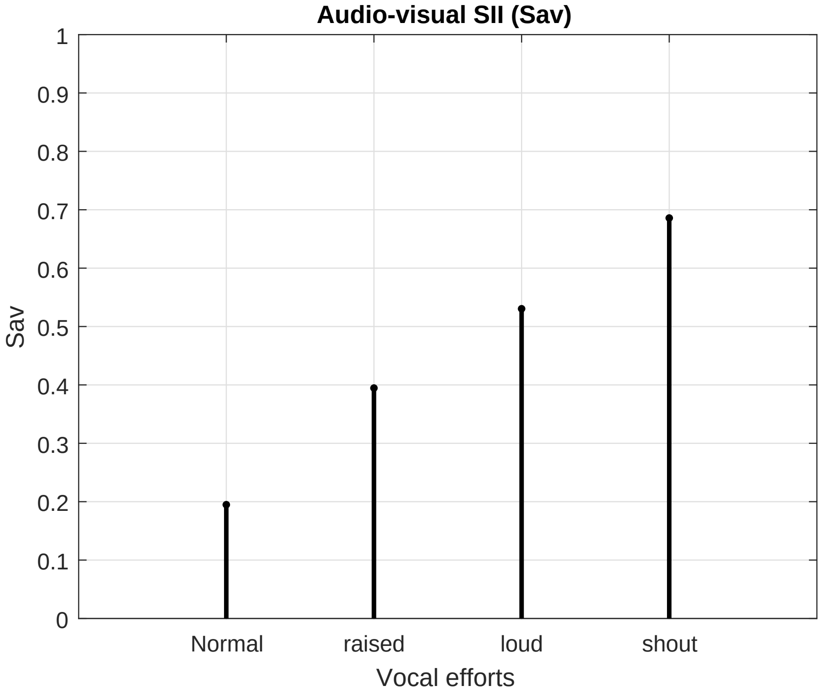


Figure 5.6: Audio-visual SII for Restaurant A

The resulting score indicated that for speech at normal conversational levels only 20% will be understood. The guests at this restaurant will need to shout to make themselves heard. In other words, guests at this restaurant can not expect to have a comfortable experience, if they want to talk to each other.

At this point another question needs discussion. This calculated SII is based on the measured noise spectrum of a restaurant that may not be 80% occupied. So the same problem arises, namely to find the spectrum of the maximum noise level, from a measurement at lower occupancy. The suggested approach here is to not go there at all. It would introduce another layer of uncertainty, another assumption based on averages. Instead, use the SII as an indicator that (hopefully) supports the dBA measurement. The SII for Restaurant A clearly shows that speech intelligibility will be a problem and it required little extra work on site. And if the restaurant actually was at 80% capacity, the SII would be highly correlated with speech intelligibility overall.

Let’s look at why this restaurant has an insufficient acoustic quality. There’s absorption in the ceiling and a reverberation time between 0.5 and 0.7 S in the bands between 250-1000 Hz. This is within the recommended range for restaurants. And still, at approximately 65% occupancy, the measured background noise of 73.5 dBA is already above the recommended range. Ambient noise levels will rise even higher as more guests arrive. Roughly speaking, 12 more people could arrive before 80% occupancy is reached. This clearly shows that a short reverberation time is not sufficient to guarantee a comfortable experience. The number of simultaneous talkers is equally important. There could simply be too many tables in the room.

In the case of Restaurant A, it is not necessary to take any more steps. The ambient noise level is already too high for comfortable communication and that's all a potential guest needs to know. If however, the measured ambient noise level had been within the accepted range, further analysis would be needed. The already mentioned occupancy is one factor. Another factor is the cross table distance. The sum of these factors leads to a more rapid escalation of noise than the expected 6 dB per doubling of guests. As a consequence, the measured $L_{N,A}$ should be corrected even more. I.e. if the restaurant was at 40% capacity, adding the logical 6 dB (from the Lombard effect) will not be enough.

In this particular case, the measured ambient noise level was already too high for verbal communication so the corrections would not change the outcome of the evaluation. However, Restaurant A illustrates the importance of correctly identifying the problem. Especially if improvements are to be made. Then the acoustician must answer the more detailed questions. Are the acoustic properties of the restaurant an issue, or is it simply a problem of having too many guests in there? I.e. what is the acoustic capacity of the room? These can not be answered from a single measurement. It can only assess the expected quality of the experience, for the guests.

The important message to the owners of Restaurant A would be to reduce the number of tables. There is a limit to how short the reverberation time can be before it reduces privacy for the speakers. Once the room has been reasonably treated with absorptive materials, the next step is to reduce the number of seats. This would increase the distance between tables and decrease the number of simultaneous talkers. To determine how big a reduction requires more information than the single measurement of $L_{N,A}$ can give (Reverberation time and Room Volume). And again, this is not the goal of the method proposed here.

Since the test case included a more complete measurement of the room (V and T), one final ”check” can be made. By calculating the Schroeder frequency, the argument for the diffuse field assumption can be strengthened. Inserting V and T into equation 3.5 from page 10 gives

$$f_{Schroeder} \geq 2000 \cdot \sqrt{\frac{0.6}{765}} \geq 57Hz \tag{5.1}$$

The frequency is, as expected, below the are of interest in restaurants (the vocal range).

5.8 Uncertainties with the single measurement method

By necessity, any attempt to predict noise levels from groups of people is based on statistical averages. This introduces many uncertainties. One of them is the Lombard effect itself. As shown in section 3.3.1 on page 13, the increase in speech level is not constant. It varies throughout a person's vocal range. The shape of the curve, the range between max/min levels and the actual max/min values also vary significantly from person to person (Hodgson et al., 2007). It is still not known how much of the increase in vocal intensity is due to an automatic reflex (Lombard effect) or due to a learned response to listener's needs (Jr. et al., 1989). But since the Lombard effect is at least in part an attempt to maintain a constant SNR (van Heusden et al., 1979) *for the listener* it suggests that people with low speech output will raise their vocal level even more, to ”keep up” with the rising ambient noise . As a consequence, for the ”low talkers” the quality of the dining experience will be reduced at lower noise levels than for the people with loud voices. They need to exert more effort to be heard by the listener. This variability between people is relevant since the goal of this method is to predict an individual's quality of the experience, based on ambient noise level. It introduces an uncertainty with regards to where to draw the line between good and bad acoustics. The only choice here is to follow the ISO 9921 recommendations, since little data is available elsewhere.

The existing work focuses on prediction of noise levels, from acoustic variables such as ambient noise level, reverberation time, absorption and the Lombard effect. In a typical EE, noise from other talkers is the main sound source so the research naturally focuses on correctly accounting for the Lombard effect. The common model is ”under these known conditions(RT,A,V) , N number of talkers will produce an ambient noise level of $L_{N,A}$, and the talkers will use a vocal effort of $L_{N,S}$” . These predictions can then be used for classification. Examples are speech intelligibility or acoustic capacity. A conclusion in an estimated case may be ”In these conditions, speech intelligibility will be poor if the number of guests exceed N” . But can this model be used ”in reverse”? If the only known variable is ambient noise level, can the same classification be certain? If the goal is to determine how ”good” or ”bad” the acoustics are in an existing EE, what is needed? To truly answer this question in future work, a large survey of many different restaurants and groups is needed. It does put one large asterisk next to the entire method proposed here. However, the same uncertainty would be there for all uses of the ISO9921 standard. So, all of this is written under the assumption that ISO 9921 recommendations are accurate.

Another uncertainty is the importance of *visual cues*. The noise level in most EEs are certainly higher than a classroom, auditorium or conference room. But the talkers and listeners are in closer proximity to each other, so visual input plays a bigger role. The Speech Intelligibility Index has a section on this topic and includes an equation for including visual cues (3.11 on page 19).

The speech content is also likely to be more familiar to the listener, in terms of syntax and semantics. The various speech tests used to determine speech intelligibility range from nonsense syllables to normal sentences. But in a dining situation it's probable that the speakers and listeners are familiar with each other's speech patterns. Both the content and how it is expressed will likely aid the listener's comprehension to some degree. This opens the possibility that a lower SNR can be sufficient for an enjoyable experience.

What a person expects from an experience will greatly influence how it is perceived, and thus the level of satisfaction. Guests at a 5-star restaurant will have higher expectations than those at a burger joint. In EEs, the background noise level functions as a masker that gives needed privacy. It also creates energy and sets a mood. One way to account for this factor is to include the age of the guests. Restaurants at the higher end of the price scale tend to have older guests who are more conscious of noise. By adjusting for age related hearing loss, one is essentially putting higher demands on the more expensive restaurants. The resulting SII will give a better indication of guests' satisfaction. Customers at a burger joint will not have the same noise sensitivity as those in a 5-star restaurant.

5.9 Calculation Example

After the acoustician has decided how to assess the restaurant, i.e. several assessments for separate areas or if one is sufficient, a calculation could look like this:
Step 1: The ambient noise level is measured to be $L_{N,A} = 67\ dBA$
Step 2: The restaurant is 40% occupied. Add 3 dBA to the measured value.
Step 3: The average distance across table is 1.3 meters. Add approx 1.5 dBA
The estimated ambient noise level is $L_{N,A} = 71.5\ dBA$ which gives it a rating of ”bad” . If the acoustician had not factored in the cross table distance, the ambient noise level would be 70 dB, and the rating would be ”sufficient” . Likewise, if the occupancy had not been factored in, the rating would end in the wrong category.

Chapter 6

Conclusion

The goal of this work has been to develop a method for evaluating the acoustic conditions in restaurants, while raising societal awareness on the topic. A two-part strategy was chosen: (1) The evaluation is focused on predicting to quality of the dining experience, from the guests' perspective. (2) The evaluation method should be simple enough to be implemented on a large scale. In accordance with (1), the method aims to provide information to the potential guests about what comfort level to expect at the restaurant. From (2), the method requires only a single measurement and needs no participation from the restaurant's owner or staff. The measurement consists of the A-weighted ambient noise level (in occupied state) as well as the 1/3-octave frequency spectrum. The spectrum is used as the input for calculating the SII (Speech Intelligibility Index).

The ease and quality of verbal communication was chosen as the objective measure of quality. The intelligibility of speech in noise is commonly evaluated with the SNR (speech-to-noise ratio) . By measuring the A-weighted ambient noise level in a restaurant, while occupied, and comparing the noise level to a statistical speech level, the SNR can be estimated. The SNR is used to classify the expected quality of the dining experience. One specific measure of verbal communication quality is the Speech Intelligibility Index (SII). The SII is a useful supplement to the dBA value and may in some cases uncover a specific frequency-problem that the A-weighting missed. Since the necessary input to the calculation can be obtained at the same time as the dBA, it does not make the measurement process any more complex.

The feedback nature of a multitalker situation (sometimes called the Cocktail party effect) is factored in to allow for measurements at lower occupancy rates. Existing models of the Lombard effects allows for prediction of the ambient noise level, as more and more guests arrive.

Cross table distance is also factored in for more accurate prediction of the SNR. As sound propagates it loses level. Larger tables will result in poorer SNR at the listener's ear. However, since part of the Lombard effect is a learned response to the listener's need, the talkers will compensate to some degree. They will pick up on visual cues and speak louder, to maintain a comfortable SNR for the listener. To factor this in, the distance corrections should be less than the usual 1/r (inverse distance law) of sound propagation.

The resulting rating is based on the recommendations for in ISO 9921. To make the assessment useful to the general public, the rating consist of only 3 classes: ”Good”, ”Okay” and ”Bad”, each with a corresponding ”smiley” symbol. The purpose of this simplicity is to make it as easy as possible to communicate the rating. The hope is that a rating will validate people’s subjective experiences at a restaurant. If guests can easily confirm that a noise level they just experienced is in fact *too loud*, they might start to change their habits. If this happens, the restaurants with bad acoustics will begin to suffer the consequences.

Bibliography

ANSI-S3.5, 1997. ANSI S3.5-1997. American National Standards Institute, 1st Edition.

Bitner, M. J., 1992. Servicescapes: The impact of physical surroundings on customers and employees. Journal of Marketing 56 (2), 57–71.

Bradley, J. S., 1986. Predictors of speech intelligibility in rooms. The Journal of the Acoustical Society of America 80 (3), 837–845.

Chen, X., Kang, J., 2017. Acoustic comfort in large dining spaces. Applied Acoustics 115, 166 – 172.

Hodgson, M., 1996. When is diffuse-field theory applicable? Applied Acoustics 49 (3), 197 – 207.

Hodgson, M., Steininger, G., Razavi, Z., 2007. Measurement and prediction of speech and noise levels and the lombard effect in eating establishments. The Journal of the Acoustical Society of America 121 (4), 2023–2033.

ISO, 2003. ISO 9921:2003 Ergonomics - Assessment of speech communication. International Organization for Standardization, 1st Edition.

Jr., H. L. P., Siegel, G. M., Fox, P. W., Garber, S. R., Kearney, J. K., 1989. Inhibiting the lombard effect. The Journal of the Acoustical Society of America 85 (2), 894–900.

Kinsler, L., Frey, A., Coppens, A., Sanders, J., 2000. Fundamentals Of Acoustics. Wiley.

Korn, T. S., 1954. Effect of psychological feedback on conversational noise reduction in rooms. The Journal of the Acoustical Society of America 26 (5), 793–794.

Lazarus, H., 1986. Prediction of verbal communication is noise a review: Part 1. Applied Acoustics 19 (6), 439 – 464.

Lazarus, H., 1987. Prediction of verbal communication in noisea development of generalized sil curves and the quality of communication (part 2). Applied Acoustics 20 (4), 245 – 261.

Lindborg, P., 2015. Psychoacoustic, physical, and perceptual features of restaurants: A field survey in singapore. Applied Acoustics 92, 47 – 60.

Long, M., 2005. Architectural Acoustics. Elsevier.

Novak, C. C., Lopa, J. L., Novak, R. E., 2010. Effects of sound pressure levels and sensitivity to noise on mood and behavioral intent in a controlled fine dining restaurant environment. Journal of Culinary Science & Technology 8 (4), 191–218.

Pickett, J. M., 1958. Limits of direct speech communication in noise. The Journal of the Acoustical Society of America 30 (4), 278–281.

Raab, C., Zemke, D. M. V., Hertzman, J. L., Singh, D., 2013. Restaurant customers perceptions of noise and their satisfaction and loyalty behaviors. International Journal of Hospitality & Tourism Administration 14 (4), 398–414.

Rhebergen, K. S., Versfeld, N. J., 2005. A speech intelligibility index-based approach to predict the speech reception threshold for sentences in fluctuating noise for normal-hearing listeners. The Journal of the Acoustical Society of America 117 (4), 2181–2192.

Rindel, J. H., 2010. Verbal communication and noise in eating establishments. Applied Acoustics 71 (12), 1156 – 1161.

van Heusden, E., Plomp, R., Pols, L., 1979. Effect of ambient noise on the vocal output and the preferred listening level of conversational speech. Applied Acoustics 12 (1), 31 – 43.

Webster, J. C., Klumpp, R. G., 1962. Effects of ambient noise and nearby talkers on a facetoface communication task. The Journal of the Acoustical Society of America 34 (7), 936–941.

Zemke, D. M., Hertzman, J. L., Raab, C., Singh, D., 2011. A little more noise and a little less conversation? ambient noise in restaurants. Journal of Foodservice Business Research 14 (3), 256–271.

Appendix

6.1 Matlab code

The following code is available from ASA, the developers of ANSI 3.5
Author: Hannes Müsch (with additions from Pat Zurek)
Available at: http://www.sii.to/html/programs.html

6.1.1 Matlab implementation of section 4

Section 4 of the ANSI 3.5 contains the core SII calculation.

Parameters are passed to the procedure through pairs of ”identifier” and corresponding ”argument”. Identifiers are always strings. Possible identifiers are:

- ’E’ Equivalent Speech Spectrum Level (Section 3.6 in the standard)
- ’N’ Equivalent Noise Spectrum Level (Section 3.15 in the standard)
- ’T’ Equivalent Hearing Threshold Level [dBHL] (Section 3.23 in the standard)
- ’I’ Band Importance function (Section 3.1 in the standard)

Except for ’E’, which must be specified, all parameters are optional. If an identifier is not specified a default value will be used. Pairs of identifier and argument can occur in any order. However, if an identifier is listed, it must be followed immediately by its argument.

Possible arguments for the identifiers are:

- Arguments for ’E’: A row or column vector with 18 numbers stating the Equivalent Speech Spectrum Levels in dB in bands 1 through 18.
- Arguments for ’N’: A row or column vector with 18 numbers stating the Equivalent Noise Spectrum Levels in dB in bands 1 through 18.
  If this identifier is omitted, a default Equivalent Noise Spectrum Level of -50 dB is assumed in all 18 bands (see note in Section 4.2).

- Arguments for 'T': A row or column vector with 18 numbers stating the Equivalent Hearing Threshold Levels in dBHL in bands 1 through 18.
  If this identifier is omitted, a default Equivalent Hearing Threshold Level of 0 dBHL is assumed in all 18 bands .
- Arguments for 'I': A scalar having a value of either 1, 2, 3, 4, 5, 6, or 7. The Band-importance functions associated with each scalar are:
  1. Average speech as specified in Table 3 (DEFAULT)
  2. various nonsense syllable tests where most English phonemes occur equally often (as specified in Table B.2)
  3. CID-22 (as specified in Table B.2)
  4. NU6 (as specified in Table B.2)
  5. Diagnostic Rhyme test (as specified in Table B.2)
  6. short passages of easy reading material (as specified in Table B.2)
  7. SPIN (as specified in Table B.2)

The function returns the SII of the specified listening conditions, which is a value in the interval [0, 1].