Improved speech communication in a car

Examensarbete utfört i Kommunikationssystem och Reglerteknik
vid Linköpings tekniska högskola
av
Mårten Nygren

LiTH-ISY-EX-3397-2003
Handledare: Niklas Österström, Svante Björklund
Examinator: Fredrik Gustafsson
Linköping den 5 juni 2003
Abstract
In modern cars a lot of effort is put on reducing the background noise level. Despite these efforts it is often difficult for persons in the rear seat(s) to hear the person(s) in the front seat(s), due to the background noise in combination with the geometry and the acoustics properties of the passenger compartment.

The aim of this thesis was to implement a speech enhancement system to increase the audibility between the driver and the rear passenger(s). The speech enhancement system should not affect the directionality of the speech nor increase the background noise level.

A speech enhancement system has been implemented on a DSP-system in a test car. A microphone was placed in front of the driver to collect his/her speech. The microphone signal was highpass filtered to remove the main part of the background noise. The signal was delayed before it was sent out in the car standard rear loudspeakers. The delay made the speech from the driver reach the rear passenger before the sound from the rear loudspeakers. This delay was enough to get the right directionality of the sound, i.e. making the speech sound as if it only came from the driver instead of the rear loudspeakers.

In the thesis other methods to reduce background noise and get directionality of the sound were evaluated, but not implemented in the test car.

The evaluations of the system showed that the audibility was increased. At the same time the background noise level was not noticeable increased.
PREFACE

This work presented in this Master thesis is the final part of my Master of Science degree in applied physics and electrical engineering at Linköpings Universitet. The thesis has been performed at A2acoustics AB in Linköping during spring 2003.

A number of people made this thesis possible by contributing in different ways, both professionally and personally:

Professor Fredrik Gustavsson and doctoral student Svante Björklund, University of Linköping, for their valuable discussions during this thesis

Niklas Österström, for supervision on theoretical and practical issues as well as on focus and delimitations of the study.

I am grateful to my opponent Stina Klomark, for help and guidance with the rapport.

I would like to thank the staff at A2acoustics, Urban Emborg, Niklas Österström, Fredrik Samuelsson, Mats Gustavsson and Arni Ingvarsson, for making me feel welcome during my work there.

Finally I would like to express my appreciations to Tea Folkeson for her support and encouragement during my work on this thesis.

Linköping, June 2003

Mårten Nygren
ABSTRACT

In modern cars a lot of effort is put on reducing the background noise level. Despite these efforts it is often difficult for persons in the rear seat(s) to hear the person(s) in the front seat(s), due to the background noise in combination with the geometry and the acoustics properties of the passenger compartment.

The aim of this thesis was to implement a speech enhancement system to increase the audibility between the driver and the rear passenger(s). The speech enhancement system should not affect the directionality of the speech nor increase the background noise level.

A speech enhancement system has been implemented on a DSP-system in a test car. A microphone was placed in front of the driver to collect his/her speech. The microphone signal was highpass filtered to remove the main part of the background noise. The signal was delayed before it was sent out in the car standard rear loudspeakers. The delay made the speech from the driver reach the rear passenger before the sound from the rear loudspeakers. This delay was enough to get the right directionality of the sound, i.e. making the speech sound as if it only came from the driver instead of the rear loudspeakers.

In the thesis other methods to reduce background noise and get directionality of the sound were evaluated, but not implemented in the test car.

The evaluations of the system showed that the audibility was increased. At the same time the background noise level was not noticeable increased.

The work has been performed at A2 Acoustics AB in Linköping, during spring 2003.
NOTATION

Symbols

s(t) Sound time history signal
r(t) Reference time history signal
y(t) Loudspeaker time history signal
n(t) Noise time history signal
a(t) Microphone array time history signal
e(t) Error signal
E[] Expected value
p Pressure [Pa]
P Power [W]
I Intensity [W/m²]
f Frequency [Hz]
f₀ One fixed frequency
c Speed of sound
λ Wavelength (λ=c/f)
* Convolution
j \sqrt{-1}
d Distance between microphones
t_d Time delay between microphones

Abbreviations

dB Decibel
DSP Digital Signal Processor
HRTF Head Related Transfer Function
IAD Inter-aural Amplitude Difference
ITD Inter-aural Time Delay
LMS Least Mean Square
RLS Recursive Least Squares
SPL Sound Pressure Level
# TABLE OF CONTENTS

1. INTRODUCTION.................................................................................................1  
   1.1. Background.........................................................................................1  
   1.2. Task and limitations.........................................................................1  
   1.3. Method .............................................................................................1  
   1.4. Report structure...............................................................................2  
   1.5. A2 Acoustics AB Presentation.........................................................2  

2. THEORY ..............................................................................................................3  
   2.1. Sound..................................................................................................3  
      2.1.1. Sound intensity and transmission .............................................3  
      2.1.2. Room acoustics.........................................................................4  
      2.1.3. Acoustic Echo...........................................................................5  
   2.2. Hearing..................................................................................................5  
      2.2.1. Speech and hearing frequencies ...............................................5  
      2.2.2. Inter-aural Amplitude Difference..........................................6  
      2.2.3. Inter-aural Time Delay.............................................................7  
      2.2.4. Head Related Transfer Function..........................................7  
   2.3. Psychological hearing...........................................................................8  
      2.3.1. The law of the first wavefront.................................................8  
   2.4. Audibility tests...................................................................................8  
   2.5. Noise figure in car.............................................................................9  
   2.6. Filters ................................................................................................9  
   2.7. Adaptive signal processing...............................................................13  
      2.7.1. Least-Mean-Square.................................................................13  
      2.7.2. Coherence...............................................................................14  
   2.8. Microphone arrays...........................................................................15  
      2.8.1. Linear arrays...........................................................................15  
      2.8.2. Two-dimensional arrays.........................................................19  

3. TEST EQUIPMENT............................................................................................25  
   3.1. Multichannel data acquisition tools...............................................25  
   3.2. Test car.............................................................................................25  
   3.3. Loudspeakers...................................................................................25  
   3.4. Microphones.....................................................................................26  
      3.4.1. Condenser and electret microphones....................................26  
   3.5. Test equipment for the tests.............................................................27  

4. METHOD............................................................................................................28  
   4.1. Noise reduction................................................................................28  
      4.1.1. Highpass filter........................................................................28  
      4.1.2. Noise reduction using LMS.....................................................31  
      4.1.3. Microphone array....................................................................32  
      4.1.4. Conclusions............................................................................36
4.2. **Hearing direction test** ................................................................. 37
   4.2.1. Hearing direction test using time delay (subjective) ............... 37
   4.2.2. Result of hearing direction test using time delay (subjective) ................................................................................................................. 38
   4.2.3. Hearing direction test using transfer function (subjective) .... 39
   4.2.4. Result hearing direction test using transfer function (subjective) ................................................................................................................. 40
   4.2.5. Conclusions ........................................................................... 40

4.3. **Test system** ............................................................................ 41
   4.3.1. General .................................................................................. 41
   4.3.2. Result test system .................................................................. 41
   4.3.3. Conclusions ........................................................................... 42

4.4. **Hearing test** ........................................................................... 42
   4.4.1. Audibility test ....................................................................... 42
   4.4.2. Results .................................................................................. 43
   4.4.3. Conclusions ........................................................................... 43

5. **SUMMARY OF RESULTS** .......................................................... 44

6. **CONCLUSIONS** ........................................................................ 45

7. **RECOMMENDATIONS** ............................................................. 46

8. **REFERENCES** ........................................................................... 47

APPENDIX A Test Linear microphone array ........................................ 48
APPENDIX B Test sheet for the hearing test ......................................... 51
Chapter 1 Introduction

1. INTRODUCTION

1.1. Background

In modern cars a lot of effort is put on reducing the background noise, i.e. engine noise, wind noise, road noise etc. Despite these efforts it is many times difficult for the rear passengers to hear the driver and the co-driver.

Enhancing the communication between front and rear seat is consequently of great interest. In this thesis work a real time speech enhancement system between front and rear seats is developed. Methods for background noise suppression in the microphone signal as well as methods for directionality of speech are tested and evaluated.

1.2. Task and limitations

The task was to investigate the possibility of enhancing the communication between the passengers inside a car. The aim was to implement a communication enhancement system in a car so that the technology could be evaluated and demonstrated. The standard loudspeakers in the car should be used.

The task was limited to communication from the driver position to the backseat of the car.

1.3. Method

To get ideas on how to approach the task a literature review on e.g. communication systems (i.e. telephone conference systems) and the human hearing mechanism were performed.

The task was then divided into three parts:

1. The speech microphone signal(s). The background noise in the signal should be reduced.

2. Speech direction. The persons in the backseat of the car should not understand that the speech is amplified and sent out in loudspeakers.
3. Implementation of the communication enhancement system, using the results of part one and two.

1.4. Report structure

The report is divided into nine sections, which are listed below:

Section 2, ‘Theory’, including the theory about different ways of tricking the brain where the sound comes from and ways to reduce the background noise in a microphone signal.

Section 3 ‘Test equipment’, the equipment that have been used is introduced.

Section 4 ‘Method’, all tests that were performed in the car are presented.

Section 5 ‘Summary of results’, the results in section 4 are summarized.

Section 6-7 ‘Conclusions’ and ‘Recommendations’.

Section 8 ‘References’.

The last part, which is not numbered, contains appendix.

1.5. A2 Acoustics AB Presentation

A2 Acoustics AB is a spin-off company from Saab AB. The company provides efficient and innovative solutions in the field of noise and vibration for customers in the areas of Aerospace and Automotive. The services include all stages of development from idea to implementation into customer products. Intellectual property protection through patents is an integral part of the strategy.

The company has extensive experience in applying active and passive noise control measures. A number of advanced measurement and analysis methods for sound and vibration have been developed. A2 Acoustics AB also participates in a number of national and international research programs.
2. THEORY

In this chapter some basic theories on acoustics relations and fundamentals of hearing are introduced. Background noise in a driving car is described as well as three different methods to remove background noise from a microphone signal (i.e. using filter, adaptive signal processing and microphone array). Initially the concept of sound is explained.

2.1. Sound

2.1.1. Sound intensity and transmission

Sound constitutes of pressure fluctuations in the air, water or some other compressible fluid. To the left in Figure 2-1 a tuning fork is shown. Vibrations in the fork create high and low pressure fluctuations in the air that is spread out in the room. To the right of Figure 2-1 low pressure is represented with negative amplitude and high pressure with positive amplitude.

![Diagram of sound waves and tuning fork](image)

Figure 2-1. To the left a tuning fork and to the right the air pressure fluctuations that the fork creates.

The amplitude of the pressure is usually given in decibels and then called Sound Pressure Level (SPL). The formula to calculate SPL from the pressure ($p$ [Pa]) is given in equation (2-1), where $p_0$ is the reference sound pressure ($p_0 = 2 \cdot 10^{-5}$ [Pa]).

$$SPL = 20 \log_{10} \left( \frac{|p|}{p_0} \right)$$  \hspace{1cm} (2-1)
The intensity \( I \) is proportional to the square of pressure. The pressure fluctuations can be measured with an instrument (e.g. a microphone). The power \( P \) of the sound is equal to the intensity \( I \) multiplied with the area \( A \) that the intensity is spread over (i.e. \( P = I \times A \)). The intensity of the sound varies with the distance \( r \) from the source. An ideal type of source is an isotropic loudspeaker, i.e. a loudspeaker that transmits sound equal in all directions. If the loudspeaker is in free space the area normal to the direction of sound propagation is \( 4\pi r^2 \). This implies that the intensity of sound decreases with the square of the distance from the source. More information about sound intensity and transmission is found in references [1, 9, 11].

2.1.2. Room acoustics

When an acoustic wave traveling in one medium encounters another medium, a part of it is reflected and another part is absorbed. A measure on materials ability to absorb an acoustic wave is the absorption coefficient. The definition of the absorption coefficient is absorbed wave energy divided with incident wave energy, i.e.

\[
\text{Absorption coefficient} = \left( \frac{\text{absorbed wave energy}}{\text{incident wave energy}} \right)
\]  

(2-2)

A room where no reflections appear (i.e. the absorption coefficient is equal to one) is called an anechoic room.

Ordinary rooms are not anechoic. This implies that a listener receives reflections as well as direct sound. If the reflected wave is delayed more than 100 ms compared to the direct wave, the listener perceives the wave as unique. The speed of sound in air is approximately 340 m/s, implying that the path difference needs to be more than \( 340 \times 0.1 = 34 \text{ m} \) for unique echo-effects to occur because of reflected sound waves on the walls, see Figure 2-2.

![Figure 2-2. A path difference of 34 m produces an unique echo effect.](image)
A way of describing a room’s acoustic property is to calculate or measure the impulse response from a particular source position to a receiver position. The impulse response describes how the transmitted sound is transformed from source to receiver.

### 2.1.3. Acoustic Echo

If the loudspeaker is in the same room as the microphone there is a risk for acoustic echo. Acoustic echo means that the microphone picks up the sound transmitted by the loudspeaker and sends it back as an echo, see Figure 2-3. If the delay in the echo is small, it will not be noticed. If the attenuation between the loudspeaker and the microphone is high, the acoustic echo effect will be negligible.

![Figure 2-3. A situation were acoustic echo occurs, if the signal to the microphone from is strong enough.](image)

### 2.2. Hearing

The brain uses several methods to determine where the sound comes from. In the following sections some of these methods are discussed. In addition a short description of human hearing and speech is given. More information about hearing is found in references [1, 9, 11].

#### 2.2.1. Speech and hearing frequencies

The human ear can respond to sound in the frequency region between 20 – 20 000 Hz. The “speech banana” in Figure 2-4 describes the dominant frequency regions of speech and sound pressure level of normal conversation speech. Different speech sounds have different intensity maximums because the nose, palate and mouth filter the frequencies in different ways.
Figure 2-4. The speech banana, Region A contains fundamental tones, B vowels, C ringing consonants and D voiceless consonants.

The most important frequency region for speech detection is from 500 to 4000 Hz according to SAME [2]. In that range most important information exist for speech intelligibility. In mobile communication the frequency range in use is 300 to 3400 Hz. Of course it would be better if more frequencies were used, but there may be a lot of background noise in these regions and in addition it can be hard to sample with high enough sampling frequency.

2.2.2. Inter-aural Amplitude Difference

An acoustic wave reaches the ears with different amplitudes if the source is not straight in front of or behind the head. The amplitude difference depends partly on the difference in distance between the ears, but mostly because the fact of that the head more or less shadows the sound for one ear. This does, however, not imply that the ear in shadow does not get any information. Sound is delivered to that ear through diffraction around the head. Diffraction occurs mainly for low frequencies with wavelengths \( \lambda = c/f \) longer than the path around the head. The total amplitude difference between the ears is called Inter-aural Amplitude Difference (IAD) (see Figure 2-5).
2.2.3. **Inter-aural Time Delay**

Due to the possible difference in distance from the source to the ears there will also be a time delay between the two ears. This time delay is called Inter-aural Time Delay (ITD) see Figure 2-6.

![ITD Diagram](image1.png)

**Figure 2-6.** Difference in path time creates ITD.

2.2.4. **Head Related Transfer Function**

The sound perceived by a person has, as mentioned in section 2.2.2 and section 2.2.3, different amplitude and time appearance depending on the placement of the source with respect to the person. The transfer function describing the path from the eardrum to a certain position outside the ear is called Head Related Transfer Function (HRTF). The HRTF include both ITD and IAD.
2.3. Psychological hearing

It is possible to trick the ear and the brain into believing that the sound comes from other positions than from the actual source. This may be done by spectral shaping of the original sound signal using HRTF:s and time delays and taking into account crosstalk\(^1\) effects as well as room reflections. The result of the sound processing is commonly called 2D or 3D-sound and is used in for example television and computers. Gelfand [9] presents more theory about psychological hearing.

2.3.1. The law of the first wavefront

The law of the first wavefront (also called the Haas effect) states that if several sound sources are transmitting the same sound at the same time, the brain uses the nearest one for directionality information, see Figure 2-7. This is a strong effect. Even if the sound from the loudspeaker that is located longer from the listener is up to about 8 dB stronger than the closer source, the listener believes that the sound comes from the nearest source.

![Figure 2-7. d_1>d_2, the person believes that all sound comes from loudspeaker 2.](image)

2.4. Audibility tests

In the thesis the goal was to improve the speech communication in a car. In order to evaluate the performance of a communication system and to get a measure of the performance, an audibility test can be performed.

In an audibility test the capability for a person to understand what is presented is tested. The result depends on which test method that is used.

\(^1\) When using e.g. two loudspeakers, the left ear hears a little portion of the right-channel speaker. This is called crosstalk.
It is easier to catch sentences than to catch single words because the brain can use the extra information to repair lost parts of the sentence.

There are two ways to test the audibility of sentences, words, or numbers: in closed or open tests. In a closed test the person chooses between limited numbers of sentences, words, or numbers. In an open test the test person does not know which sentences, words, or numbers that will be presented. The tests may be performed with or without background noise.

2.5. Noise figure in car

In a car there are mainly four sound experiences:

1. Road noise
2. Wind noise
3. Engine noise
4. Other sounds in the car compartment (e.g. music and speech)

When traveling at different speeds, under different road conditions, or with different accelerations, a car emits different noise figures.

The totally noise figure for the first three noise components is mainly in the low frequency region.

2.6. Filters

When microphones record a person talking in a driving car, the microphone signal will contain a combination of talk and background noise (see section 2.5). To reduce the background noise in the signals, digital filters can be applied. Theory about filters and digital signals are found in Söderkvist [7].

The definition of a filter is a transformation from an incoming sequence of numbers to an outgoing sequence. Figure 2-8 shows an example of an incoming sequence \( x[n] \) that is transformed to \( y[n] \) by the filter \( h[n] \) (\( n \) represents discrete time sample).
The time signal may also be presented in the frequency domain, using the fourier transform, equation (2-3).

\[ H(f) = \int_{-\infty}^{\infty} h(t) e^{-j2\pi ft} dt \] (2-3)

To transform from the frequency domain to the time domain, the inverse fourier transform should be used, equation (2-4).

\[ h(t) = \int_{-\infty}^{\infty} H(f) e^{j2\pi ft} df \] (2-4)

There are four fundamental different types of frequency filters:

- Lowpass
- Highpass
- Bandpass
- Bandstop

Figure 2-9 shows an ideal bandpass filter with the cutoff frequencies \( f_1 \) and \( f_2 \).

\[ \text{Amplitude} \]
\[ 0 \quad 1 \]
\[ \text{Frequency} \]
\[ f_1 \quad f_2 \]

Figure 2-9. Bandpass filter with cutoff frequencies \( f_1 \) and \( f_2 \).
Equation (2-5) describes the ideal bandpass filter in Figure 2-9 in the frequency domain. Observe that both positive and negative frequencies are used to describe the filter.

\[
H_{\text{IDEAL}}(f) = \begin{cases} 
1 & -f_2 \leq f \leq -f_1 \\
1 & f_1 \leq f \leq f_2 \\
0 & \text{otherwise}
\end{cases} \tag{2-5}
\]

To transform the filter into the time domain, equation (2-4) should be applied on the expression in equation (2-5).

\[
h_{\text{IDEAL}}(t) = \int_{-\infty}^{\infty} H(f)e^{j2\pi ft} df = \int_{-f_2}^{-f_1} e^{j2\pi ft} df + \int_{f_1}^{f_2} e^{j2\pi ft} df \\
= \left[ \frac{e^{j2\pi ft}}{j2\pi} \right]_{-f_2}^{-f_1} + \left[ \frac{e^{j2\pi ft}}{j2\pi} \right]_{f_1}^{f_2}
\]

\[
\cos(2\pi f_2) + j\sin(2\pi f_2) - \cos(2\pi f_1) - j\sin(2\pi f_1) + \frac{2j\sin(2\pi f_2) - 2j\sin(2\pi f_1)}{j2\pi} = \frac{\sin(2\pi f_2) - \sin(2\pi f_1)}{\pi} \tag{2-6}
\]

The time domain bandpass filter in equation (2-6) is shown in Figure 2-10.

![Figure 2-10. Ideal bandpass filter in the time domain.](image.png)
The bandpass filter in Figure 2-10 is a time continuous filter with infinity extent. In a real-time computer system the filter must be causal and time discrete. Figure 2-11 shows the filter in Figure 2-10 displaced, truncated and sampled with 64 samples, making it causal and time discrete.

Figure 2-11. A causal time discrete bandpass filter.

Figure 2-12 shows the discrete and causal bandpass filter in the frequency domain with normalized cutoff frequencies $f_1=0.1$ and $f_2=0.3$. In the figure it is seen that the filter differs from the ideal filter in Figure 2-9.

Figure 2-12. The discrete causal bandpass filter in frequency domain.
2.7. Adaptive signal processing

If a noise spectrum is known and looks the same all the time it can be removed from the signal with a stationary filter. Usually, however the spectrum changes with time. This implies that the filter must change with time in order to reduce the noise. Several algorithms for implementing a time varying filter exist, for example:

- Recursive Least Squares (RLS)
- Kalman filtering
- Least Mean Square (LMS)

The algorithms are described in Widrow et al. [5] and Gustafsson et al. [3]. In the next section the LMS-algorithm is explained. The two other algorithms were not in the scope of this thesis work.

2.7.1. Least-Mean-Square

The LMS-algorithm is a closed loop algorithm. In the simplest mode the algorithm uses two inputs; a reference and an error signal, see Figure 2-13. The algorithm is formed to minimize the squared error signal by updating the adaptive filter, $h[n]$ i.e.;

$$\min_{h[n]} E\left[e[n] - e[n] \right]$$

(2-7)

This filter is changing with time since the error $e[n]$ is changing with time. The performance of the LMS-algorithm is limited by the coherence between the reference signal $(r)$ and the disturbance signal $(d)$. The coherence and the theoretical result are described in section 2.7.2.

![Figure 2-13. The structure of a feed-forward system using LMS.](image)
2.7.2.  Coherence

To achieve large noise reduction using the LMS-algorithm, the coherence ($\gamma$) between the reference signal ($r$) and the disturbance signal ($d$) must be high. The maximum theoretical noise reduction according to Kuo and Morgan [10] is given in equation (2-8).

$$\text{Max reduction} = 10 \cdot \log_{10}(1 - |\gamma_{dr}(f)|^2) \ [\text{dB}],$$  \hspace{1cm} (2-8)

where the coherence function $\gamma_{dr}(f)$ is:

$$\gamma_{dr}(f) = \frac{S_{dr}(f)}{\sqrt{S_{dd}(f)S_{rr}(f)}},$$  \hspace{1cm} (2-9)

and the cross-power spectrum $S_{dr}(f)$ is:

$$S_{dr}(f) = \int_{-\infty}^{\infty} E[d(\tau + t)r(t)]e^{-2\pi fi} \, dt$$  \hspace{1cm} (2-10)

$S_{dd}$ and $S_{rr}$ are the auto-power spectra of $d(t)$ and $r(t)$. 
2.8. Microphone arrays

2.8.1. Linear arrays

The main idea with a microphone array is to select sound coming from a particular direction and suppress the sound in all other directions.

There are several factors affecting the result of a microphone array, e.g.:

- The frequency range of the sound
- Setup and numbers of the microphones
- Source position
- Weight factors (i.e. how much of each microphone signal that is used)

In Figure 2-14 a linear array is shown. Sound of one frequency \( f_0 \) is used as source and the distance between the microphones is the same. The microphone signals are given the same weight factor.

It is assumed that the source is positioned far away from the array. This implies that sound coming from the front of the array reaches all microphones at the same time “plane-wave assumption”.

![Linear microphone array with 3 microphones and uniform distance (d). To the left the acoustic wave is coming from the front of the array and to the right with angle \( \alpha = 60^\circ \).](image)

To the left in Figure 2-14 the sound, \( s(t) = \sin(2\pi f_0 t) \) reaches all microphones at the same time. The sum of the three microphone signals, i.e. the output \( a(t) \) from the array becomes;
To the right in Figure 2-14 the sound does not reach the microphones at the same time. There is an inherent time delay between the microphones. The time delay depends on the angle $\alpha$, the distance $d$ and the speed of the sound $c$. Defining that the sound reaches the closest microphone at time $t$, then the sound reaches the next microphone $t_d$ seconds later. The time delay $t_d$ is:

$$t_d = \frac{d \cdot \sin(\alpha)}{c} \quad (2-12)$$

The sum of 3 microphones becomes:

$$a_\alpha(t) = s(t) + s(t + d \cdot \frac{\sin(\alpha)}{c}) + s(t + 2 \cdot d \cdot \frac{\sin(\alpha)}{c}) \quad (2-13)$$

In Figure 2-15 the microphone signals ($m_1$, $m_2$, $m_3$) and the array output for the linear microphone array to the right in Figure 2-14 with a specific set up of parameters is shown. The figure shows that the time delay causes suppression when adding the signals when the sound is coming $60^\circ$ from the array.

Equation (2-13) implies that $a_{\alpha=0}(t) \geq a_\alpha(t)$. Figure 2-16 shows the theoretical microphone array output amplitude for different angles and for a specific frequency.
Figure 2-16. Linear Microphone array output amplitude with N=3 microphones, d=0.07m and \( f_0=680 \text{ Hz} \).

It is also interesting to see the performance of the array for different frequencies. Equation (2-12) and (2-13), with \( s(t) = \sin(2\pi f_0 t) \), give:

\[
a_3(t) = \sin(2\pi f_0 t) + \sin(2\pi f_0 (t + t_d)) + \sin(2\pi f_0 (t + 2t_d))
\]  
(2-14)

Introduce \( \beta = 2\pi f_0 t_d \)

\[
\beta = 2\pi f_0 t_d
\]  
(2-15)

and rewrite equation (2-14):

\[
a_3(t) = \sin(2\pi f_0 t) + \sin(2\pi f_0 t) \cos(\beta) + \cos(2\pi f_0 t) \sin(\beta) + \\
\sin(2\pi f_0 t) \cos(2\beta) + \cos(2\pi f_0 t) \sin(2\beta)
\]  
(2-16)

To represent the microphone array in the frequency domain the Fourier transform as described in equation (2-3) is applied to equation (2-16).

\[
A_3(f) = (j\pi(\delta(f + f_0) - \delta(f - f_0))) \cdot (1 + \cos(\beta) + \cos(2\beta)) + \\
(\pi(\delta(f + f_0) + \delta(f - f_0))) \cdot (\sin(\beta) + \sin(2\beta))
\]  
(2-17)

where:

\[
\int_{-\infty}^{\infty} \delta(f - f_0) g(f) = g(f_0), \ g(f) \text{ is a time continues function}
\]  
(2-18)

A plot of the noise reduction as function of frequency is shown in Figure 2-17. The reduction is calculated according to equation (2-19).
\[ REDUCTION = 20 \log_{10}(\frac{\text{array}(f)}{\text{Reference microphone}(f)}) \] \hspace{1cm} (2-19)

The conditions for the array are the same as in the example in Figure 2-15 where the same array was presented in the time domain, i.e. \( d=0.07 \text{m}, \, N=3 \) and \( \alpha=60^\circ \).

![Figure 2-17. Noise reduction with respect to frequencies for \( \alpha=60^\circ \), \( N=3 \) and \( d=0.07 \text{m} \).](image)

In Figure 2-17 it is seen that large noise reduction is achieved in the frequency range between 1500 Hz to 4000 Hz while low reduction apply frequencies in the range 4500-6500 Hz. Particular at \( f_0 \approx 5600 \) Hz the suppression is zero. The time delay between the microphones is for that frequency a value where the phase difference is a multiple of \( 2\pi \) i.e.:

\[
\sin(2\pi f_0 t + \beta) = \sin(2\pi f_0 t'), \quad \text{when} \quad \beta = 2\pi n, \, n = 1, 2, 3 \ldots, \hspace{1cm} (2-20)
\]

where \( \beta \) is defined in equation (2-15).

The frequency where no suppression is achieved can be calculated using equation (2-21) derived from equation (2-15) and (2-20).

\[
f_0 = \frac{n}{t_d} = \frac{cn}{d \sin(\alpha)} \hspace{1cm} (2-21)
\]

For the parameters in Figure 2-17 this gives \( f_0 = 5609 \text{Hz} \).
2.8.2. Two-dimensional arrays

The linear array suppresses sound in one plane (i.e. for different values of the angle $\alpha$ in Figure 2-14). To achieve suppression in 3 dimensions and selection of sound straight in front of the array a two-dimensional array must be used. In Figure 2-18 there is an example of a two dimensional array. In the following text basic linear algebra will be involved. In Hackman [6] linear algebra is described.

![Two-dimensional microphone array](image)

*Figure 2-18. Incident acoustic waves towards a two-dimensional microphone array.*

To describe the incoming waves a coordinate system with x- and y-axis in the plane where the microphones are located is introduced, see Figure 2-19. Origo is placed in the same position as microphone 5 ($m_5$) in Figure 2-18. The angles are defined according to equation (2-22)

\[
0^\circ \leq \varphi \leq 90^\circ \\
0^\circ \leq \theta \leq 360^\circ
\]  

(2-22)

The definition of the angles implies that waves are considered to come only from the front of the microphones.
In Figure 2-19 the symbol \( \vec{n} \) describes the incoming acoustic waves to the microphones. The incoming wave has a given fixed length \( (r) \). The vector, \( \vec{n} \) in equation (2-23) describes the incoming wave. The normalized vector \( \hat{n} \) is calculated in equation (2-24).

\[
\vec{n} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -\sin(\theta) \cdot r \cdot \sin(\varphi) \\ -\cos(\theta) \cdot r \cdot \sin(\varphi) \\ r \cdot \cos(\varphi) \end{pmatrix}
\] (2-23)

\[
\hat{n} = \frac{\vec{n}}{||\vec{n}||} = \frac{\vec{n}}{r\sqrt{\sin^2(\theta)\sin^2(\varphi) + \cos^2(\theta)\sin^2(\varphi) + \cos^2(\varphi)}} = \frac{\vec{n}}{r\sqrt{\sin^2(\varphi) + \cos^2(\varphi)}} = \frac{\vec{n}}{r} = \begin{pmatrix} -\sin(\theta) \cdot \sin(\varphi) \\ -\cos(\theta) \cdot \sin(\varphi) \\ \cos(\varphi) \end{pmatrix}
\] (2-24)

The vector \( \hat{n} \) is orthogonal to the plane of the incoming waves. In Figure 2-20 a vector model of the distance between the plane of the incoming
waves and one of the microphones is shown. The plane is placed so that microphone \( m_5 \) is in the plane of the incoming waves.

Figure 2-20. The plane for the incoming waves and the distance \( d_i \) to microphone \( m_i \).

The distance \( d_i \) from the wave plane to the microphone \( m_i \) is the distance that the wave must travel from the reference microphone \( m_5 \) in origo to microphone \( m_i \). The distance can be positive or negative, depending on the choice of reference microphone. A negative distance means that the waves reach microphone \( m_i \) earlier than the reference microphone.

The microphone positions are, according to Figure 2-18;

\[
M = \begin{pmatrix}
    m_1 \\
    m_2 \\
    m_3 \\
    m_4 \\
    m_5 \\
\end{pmatrix} = \begin{pmatrix}
    x & y & z \\
    0 & d & 0 \\
    -d & 0 & 0 \\
    0 & -d & 0 \\
    d & 0 & 0 \\
    0 & 0 & 0 \\
\end{pmatrix}
\] (2-25)

The distance \( d_i \) is the orthogonal distance from the plane of incoming waves to microphone \( \tilde{m}_i \): 

\[
d_i = \tilde{m}_i \cdot \hat{n}
\] (2-26)

The extra distances that the waves must travel to reach each microphone thus are:
The time delay for each microphone is \( t_{d,j} = \frac{d_j}{c} \). The sum of the microphone signals becomes:
\[
a_{\theta,\varphi}(t) = s(t) + s(t + d_1/c) + s(t + d_2/c) + s(t + d_3/c) + s(t + d_4/c) + s(t + d_5/c)
\]
(2-28)

The sum is shown in Figure 2-22 for one frequency and for different angles. The distance from origo describes the angle \( \varphi \) and \( \theta \) is described as the angle to the y-axis in the first quadrant according to Figure 2-21 and equation (2-29).

\[
\begin{pmatrix}
  d_1 \\
  d_2 \\
  d_3 \\
  d_4 \\
  d_5
\end{pmatrix} =
\begin{pmatrix}
  -d \cos(\theta) \cdot \sin(\varphi) \\
  d \sin(\theta) \cdot \sin(\varphi) \\
  d \cos(\theta) \cdot \sin(\varphi) \\
  -d \sin(\theta) \cdot \sin(\varphi) \\
  0
\end{pmatrix}
\]
(2-27)

\[
\theta = \theta \quad 0^\circ \leq \theta \leq 360^\circ \\
\varphi = \sqrt{x^2 + y^2} \quad 0^\circ \leq \varphi \leq 90^\circ
\]
(2-29)
Figure 2-22. Microphone array in two-dimensions with $f_0=1500$ Hz and $N=5$

Where the white dot is at $\theta = \tan^{-1}(50/30) \approx 60^\circ$ and $\varphi = \sqrt{50^2 + (-30)^2} \approx 60^\circ$ the noise reduction is about 6 dB for frequency 1500 Hz.

Figure 2-23. Linear microphone array in 1 dimension with $f_0=1500$ Hz $d=7$cm $N=5$.

Figure 2-23 shows the summation of the microphone signals for the linear microphone array in the same type of plot as in Figure 2-22
In special cases for example when $\theta$ is fixed to $90^\circ$, the two-dimensional array could be seen as a one-dimensional array with three microphones. The weight on the centre microphone is then 3 the weight on the side microphones is 1 see Figure 2-24.

Figure 2-24. The two dimensional array with theta fixed to $90^\circ$.

$$a_{\theta=90^\circ}(t) = \sin(2\pi f t) + 3 \sin(2\pi f (t + t_d)) + \sin(2\pi f (t + 2t_d))$$ (2-30)

Equation (2-30) shows the output for the microphone array when $\theta$ is fixed to $90^\circ$ ($\theta$ is described in Figure 2-19). Figure 2-25 shows the theoretical output for the microphone array for different frequencies and a fixed angle $\varphi=60^\circ$. Compared with Figure 2-17 it is a smaller frequency region where there is large reduction.

Figure 2-25. Theoretical result for two dimensional microphone array with $\theta=90^\circ$, $\varphi=60^\circ$, $d=7$ cm and $N=5$.

More about microphone arrays can be found in Ögren [4]. In Ahlin and Zander [8] a physical description is done.
3. TEST EQUIPMENT

In the following sections a short description of the test equipment used during the tests described in chapter 4 is presented.

3.1. Multichannel data acquisition tools

To collect multichannel data, two apparatus were used:

- DAT-recorder, a Digital-Audio-Tape-recorder that was capable to record 8 channels at the same time (TEAC RD-135 T DAT-recorder).
- Leuven Measurement System, CADA-X Fourier Monitor. It is a multichannel data acquisition system for acoustical analysis.

3.2. Test car

The test car was a Saab 9-5 2.3T sedan, 1999 year's model.

3.3. Loudspeakers

A loudspeaker transforms electrical energy to acoustical energy. Of main interest in this thesis is:

- Distortion, i.e. if an electrical frequency is distorted to several acoustical frequencies or not.
- Frequency response, i.e. in what frequency ranges the loudspeaker manages to reconstruct the electrical energy to acoustical energy.

In Crocker [12] detailed information of loudspeakers are found.
3.4. Microphones

A microphone is an instrument that transforms sound pressure or particle speed to electrical voltage or current with maintained waveform. There are a number of physical properties that characterize a microphone:

- Frequency response.
- Sensitivity, the ratio between amplitude of the electrical signal and the amplitude of the sound pressure (or particle speed).
- Sensitivity to external influences (e.g. wind and temperature variations).
- Directionality, i.e. from which directions the microphone catches the sound.

There are many different types of microphones. In Crocker [12] detailed information about microphones can be found. The microphone types that were used in this thesis were so-called condenser and electret microphones and are described in the next section.

3.4.1. Condenser and electret microphones

Condenser microphones use the fact that the capacitance changes with respect to the distance between two electrical charged plates. A condenser microphone consists of one light membrane plate and one fixed plate. Pressure fluctuations put the light membrane in motion. If the plates have constant electrical charge, the motion creates changes in potential between the plates that reflect the pressure fluctuations.

Electret microphones are constructed very similar as the condenser microphone. In a condenser microphone the plates are charged with an external source. Electret microphones instead use an electret plate to create the differences in potential.
3.5. Test equipment for the tests

In Table 3-1 the test equipment used during the different tests are shown. Below the different tests are listed:

1. Highpass filter
2. LMS-algorithm
3. Microphone array
4. Hearing direction using time delay
5. Hearing direction using transfer function
6. Demonstration system
7. Audibility test

<table>
<thead>
<tr>
<th>Test Equipment</th>
<th>Used during tests</th>
</tr>
</thead>
<tbody>
<tr>
<td>DAT-recorder</td>
<td>1,2,3,4,5</td>
</tr>
<tr>
<td>Car loudspeaker</td>
<td>6,7</td>
</tr>
<tr>
<td>Loudspeakers 4 inch</td>
<td>4,5</td>
</tr>
<tr>
<td>Microphones, Norsonic. Condenser microphone</td>
<td>1,2,3,4,5</td>
</tr>
<tr>
<td>Microphones, Panasonic. Electret microphones</td>
<td>6,7</td>
</tr>
<tr>
<td>Powermax, Battery belt</td>
<td>1,2,3</td>
</tr>
<tr>
<td>Digital signal processor(Texas Instrument TMS320C32)</td>
<td>6,7</td>
</tr>
<tr>
<td>PS-72, Voltage stabilizer</td>
<td>6,7</td>
</tr>
<tr>
<td>Microphone calibrator</td>
<td>1</td>
</tr>
<tr>
<td>Leuven Measurement System</td>
<td>1,2,3,4,5</td>
</tr>
<tr>
<td>PC, equipped with MatLab</td>
<td>1,2,3,4,5</td>
</tr>
<tr>
<td>Labgruppen, Amplifier to loudspeakers</td>
<td>4,5</td>
</tr>
<tr>
<td>Sony-walkman</td>
<td>4,5,7</td>
</tr>
<tr>
<td>Norsonic amplifier type 335</td>
<td>1,2,3,4,5</td>
</tr>
</tbody>
</table>
4. METHOD

In this chapter methods to reduce the background noise and methods for speech directionality are evaluated. Also, the implementation of the communication enhancement system is described.

Finally the hearing audibility with the system on and off is evaluated.

4.1. Noise reduction

The sound in a car is a combination of background noise and speech. It is important to reduce the influence of background noise in the microphone signal. Otherwise, when sending out the signal in the loudspeakers the background noise level in the car will increase. In order to suppress the background noise from the microphone signals, three methods were tested and evaluated:

- Highpass filtering of microphone signals
- Adaptive filtering of microphone signal using the LMS algorithm
- Microphone array

4.1.1. Highpass filter

As described in section 2.5, the background noise in a driving car is dominated by low frequency noise. Figure 4-1 shows a typical noise spectrum in a car. The data was collected inside a car driving at 110 km/h.

![Figure 4-1. Noise spectrum inside in a car driving at 110 km/h.](image-url)
To reduce the main part of the background noise in the microphone signal a highpass filter can be used. A design problem is that the system sampling frequency is limited. To avoid problems with aliasing, an analogue lowpass filter before the A/D-converter should be applied. The analogue lowpass filter was however not within the scope of this thesis. Figure 4-2 shows a highpass filter in the frequency domain with cutoff frequencies 300. The filter is designed according to equation (4-1) with 64 samples in the time domain (i.e. with the MatLab\textsuperscript{®} function \texttt{fir1}). More about filter design can be found in Gustafsson et al. [3].

$$\min_{H} \int |H_{\text{IDEAL}} - H|^2 \, df$$

(4-1)

Where $H$ is the time discrete highpass filter.

![Figure 4-2. Time discrete highpass filter.](image)

Applying the highpass filter in Figure 4-2 to the noise signal from the car (Figure 4-1) results in the spectra shown in Figure 4-3.
Figure 4-3. Spectrum for a signal from a car after and before highpass filtering.

Figure 4-4 shows a microphone time signal, containing both speech and background noise recorded in a car driving in 110 km/h. The figure shows the time signal both before and after highpass filtering.

Figure 4-4. Microphone signal containing speech and background noise (110 km/h), before and after highpass filtering.

In Figure 4-4 it is noted that the speech appears when the microphone signal is filtered.
4.1.2. Noise reduction using LMS

To evaluate the possibility of using the LMS-algorithm (section 2.7.1) for background noise suppression, the coherence between seven different microphone signal positions (see Figure 4-5) were calculated (see the theory described in section 2.7.2). A good reference signal should have high coherence with the noise in the monitor microphone (i.e. the microphone that should collect the speech signal) and low coherence with the speech. The coherences were calculated on recorded signals in the driving test car. Positions 2-7 were thought of as being possibly reference positions and position 1 as monitor microphone.

![Figure 4-5. Microphone positions for the coherence test.](image)

Figure 4-6 shows the theoretical maximum noise reduction (see equation (2-8)), based on the measured coherences between the signals. It is concluded that the reduction of the background noise in the speech signal will be low if the LMS-algorithm is used. Higher coherence could possibly be achieved if other microphone positions were used. However, the reference microphone should be placed so that only the background noise in the reference microphone is coherent with the background noise in the monitor microphone (i.e. the coherence in the speech signal between the reference and the monitor microphone should be low).
4.1.3. Microphone array

Five tests have been performed to evaluate the possibility of background noise reduction using microphone arrays. The five tests were:

1. A loudspeaker fed with constant spectral level noise from 20 to 10000 Hz, placed 60° beside the microphone array, according to Figure A-1 in appendix A
2. A loudspeaker fed with constant spectral level noise from 20 to 10000 Hz, placed 60° below the microphone array, according to Figure A-2 in appendix A
3. Driving on motorway and persons inside the car were silent.
4. Driving on motorway and person in front of the array speaking (on axis person, driver).
5. Driving on motorway and person beside the array speaking (off axis person, co-passenger)

A highpass filter with cutoff frequencies 300 Hz was applied on the microphone signals before they where added to a microphone array signal for test 3 to 5. Three different setups of microphones where evaluated. Two linear microphone arrays, one with five microphones with the
uniform microphone distance 7 cm and another with three microphones with the uniform distance 25 cm. The third variant was a 2 dimensional array with five microphones with 7 cm between the microphones (see Figures A-3, A-4 and A-5 in appendix A). Each microphone was given the same weight.

4.1.3.1. Result, Test 1

In Figure 4-7 the result of test 1 is shown. The result should be compared to the theoretical results in sections 2.8.1 and 2.8.2.

The reduction is calculated according to equation (4-2).

\[
REDUCTION = 20 \log_{10} \left( \frac{|array(f)|}{|Reference \ \text{microphone}(f)|} \right)
\]  

(4-2)

The suppression for the linear microphone array with \(d=7 \text{ cm}\) is, compared to the theoretical results in Figure 2-17 (in Figure 2-17 there is only 3 microphones in the array, but the theoretical result for three and five microphones are similar), as good as it could be expected in frequency region 1000 to 4000 Hz.

Figure 4-7 shows that the linear microphone array with uniform distance \(d=25 \text{ cm}\) has small or no reduction for all frequencies.

![Figure 4-7. Noise reduction in test 1.](image)
4.1.3.2. Result, Test 2

The result of test 2 is shown in Figure 4-8.

![Graph showing noise reduction in test 2.](image)

**Figure 4-8. Noise reduction in test 2.**

When the loudspeaker is located as in test 2 it should theoretically not be any suppression for the linear microphone arrays. The reduction that occurs depends on the fact that the arrays are not placed in an anechoic room. The acoustic wave is reflected and incident on the array with a different phase than the direct wave.

The result in Figure 4-7 and Figure 4-8 for the two-dimensional array should be the same for both test 1 and 2. The graphs should be compared to Figure 2-25. The variations are caused by differences in the transfer functions between the loudspeaker and the microphone array for the two tests.
4.1.3.3. Result tests 3, 4 and 5

In test 1 and 2 the background noise to be suppressed was random noise with constant spectral density. In a driving car the background noise is, as discussed in section 2.5, dominated by low frequency noise. Figure 4-9 shows the result of test 3, 4 and 5 for the linear microphone array with uniform distance 7 cm.

![Figure 4-9. Test 3-5 in both frequency and time domain for linear array with d=7cm.](image)

All three microphone arrays work for frequencies above 1000 Hz and give about 5 dB reduction. For higher frequencies (2000 Hz) reduction of the background noise up to 10 to 15 dB are not unusual. The total reduction is however low (2.9 dB for the linear microphone array with d=7 cm.) since the low frequency noise still dominate the background noise, although the noise signal has been highpass filtered.

Better suppression in the low frequency region was achieved with the linear microphone array with 3 microphones and 25 cm distance between the microphones. A problem that occurred was the long distance between the person and the microphones. As mentioned in chapter 2.1 the intensity of sound decrease with the squared distance.
4.1.4. Conclusions

The following conclusions can be drawn about the three noise reduction methods (i.e. LMS, highpass filter, and microphone array):

- The coherence between the monitor signal and the tested reference signals were too low to suppress noise efficiently for frequencies above 300 Hz with an LMS-algorithm.

- A highpass filter with cutoff frequency 300 Hz suppressed the background noise well without affecting the speech negatively.

- With microphone arrays, large distance between the individual microphones is needed in order to suppress low frequency noise. When increasing the distance between the microphones, the distance to the speaking person becomes too large. The extra reduction that an array gives compared to one microphone is not large enough to motivate the use of several microphones instead of one microphone in this application.
4.2. **Hearing direction test**

As discussed in section 2.3, it is possible to trick the brain to believing that the sound comes from other positions than from the actual source. When sending out the driver’s speech in the car loudspeakers, it is possible that the passenger perceives the sound as if it comes from the loudspeakers. To investigate if this could be avoided, two tests using the law of the first wavefront (see section 2.3.1) and transfer functions respectively were performed.

Both tests were performed under the assumption of using the rear loudspeakers in the car as “secondary” sources for speech (the left rear loudspeaker is loudspeaker 2 in Figure 4-10). These loudspeakers are located far from the speaker, implying that the driver may not have to hear his/her own speech from the loudspeakers, which is good.

Below the details of the tests are given.

4.2.1. **Hearing direction test using time delay (subjective)**

The methodology for this test is based on the law of the first wavefront. The test was performed in the following way:

- One sentence was first recorded on a DAT-recorder.
- The recorded sentence was transformed to wave-format so that it could be treated in Matlab. Wave format is a stereo format, i.e. the samples are put in two vectors, one for the right loudspeaker and one for the left loudspeaker. When the recording was done, only one microphone was used, implying that both vectors were the same.
- One of the two vectors was time delayed.
- The sound was then written on a CD.
- The CD was played with the time-delayed signal through loudspeaker 2 and the original signal through loudspeaker 1 see Figure 4-10.

To see if the rear passenger perceived the sound as coming from the front speaker, several time delays were tested; 0, 5, 10, 15 and 20 ms. The test persons had to indicate if the sounds came from loudspeaker 1, 2 or maybe both. The test persons should also say if they heard any echo. The influence of different levels of the signals was not tested.
4.2.2. Result of hearing direction test using time delay (subjective)

The test was conducted on 5 persons. The results are listed in Table 4-1.

Table 4-1. Result hearing direction test using time delay, 1=loudspeaker 1, 2=loudspeaker 2.

<table>
<thead>
<tr>
<th>Test person</th>
<th>0 ms delay</th>
<th>5 ms delay</th>
<th>10 ms delay</th>
<th>15 ms delay</th>
<th>20 ms delay</th>
</tr>
</thead>
<tbody>
<tr>
<td>Test person 1</td>
<td>2</td>
<td>1,2</td>
<td>1</td>
<td>1</td>
<td>echo</td>
</tr>
<tr>
<td>Test person 2</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>echo</td>
<td>echo</td>
</tr>
<tr>
<td>Test person 3</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>echo</td>
<td>echo</td>
</tr>
<tr>
<td>Test person 4</td>
<td>2</td>
<td>1,2</td>
<td>1</td>
<td>1</td>
<td>echo</td>
</tr>
<tr>
<td>Test person 5</td>
<td>2</td>
<td>1,2</td>
<td>1</td>
<td>1</td>
<td>echo</td>
</tr>
</tbody>
</table>

In Table 4-1 the result shows that most of the test persons prefer a time delay of 10 ms before the other options. With no or with a short delay (5 ms) the test persons perceived that the sound came from loudspeaker 2. Too long time delay resulted in echo according to the test persons.
4.2.3. **Hearing direction test using transfer function (subjective)**

Another way of tricking the brain to believe that the sound from the rear loudspeaker comes from the driver position is to filter the speech signal with special transfer functions. In this test the same speech sentences as used in the previous test (section 4.2.1) was utilized. The signal to the rear speaker (2) was filtered, while the signal to the front speaker (1) was left unfiltered.

Two transfer functions are needed to transform the sound in the rear speaker, $h_{13}$ and $h_{L33}$. $h_{13}$ describes the transfer function from person 1 to person 3. $h_{L33}$ describes the transfer function from loudspeaker 3 to person 3.

![Figure 4-11. Transfer function between person 1 and person 3 and person 3 and loudspeaker 3.](image)

The goal was that the person in the backseat should catch the sound from loudspeaker 3 as if it came from the person in the driving seat. This should be possible to achieve if the signal to the rear speaker is filtered with a transfer function $h$ before it is sent out from loudspeaker 3. In equation (4-3) the wanted relation is described.

\[ s_1 * h_{13} = y_3 * h_{L33} * h \]  

(4-3)

Where $s_1$ describes the speech from person 1 and $y_3$ the signal to loudspeaker 3. In the test a recorded speech signal was used instead of a real person. Therefore $s_1$ is changed to $y_j$, i.e. the signal to loudspeaker 1, and $h_{13}$ to $h_{L13}$, i.e. the transfer function between loudspeaker 1 and person 3. Doing these changes and converting equation (4-3) into the frequency domain gives:

\[ Y_j H_{L13} = Y_3 H_{L33} H \]  

(4-4)

Rearranging gives:
The signal to the rear speaker should be filtered with the inverse fourier transform of $H$ to transform the sound from the rear speaker to sound as if it comes from the front speaker.

\[ H = \frac{Y_1 H_{L13}}{Y_3 H_{L33}} = \frac{H_{L13}}{H_{L33}}, \text{ since } Y_1 = Y_3 \]  \hspace{1cm} (4-5)

4.2.4. Result hearing direction test using transfer function (subjective)

The directionality of the sound from the rear speaker was not markedly changed after filtering the signal. When measuring the transfer functions $h_{L13}$ and $h_{L33}$, a microphone placed close to the ear canal entrance was used. It was suspected that the microphones should be placed further in the ear canal in order to include the acoustic property of the ear and to catch the right directionality of the sound. Tests with ear canal probe microphones were however not performed within the scope of this thesis work.

4.2.5. Conclusions

Two different methods for tricking the brain to believing that the sound transmitted through the rear loudspeaker instead came from the front loudspeaker (driver seat) were evaluated. The following conclusions are drawn:

- A 10 ms time delay on the signal to the rear speaker worked to trick the test person to believe that all sound came from the speaker in the front seat.

- Filtering the signal to the rear speaker with transfer functions did not work. The result would probably be better if the transfer function had been measured using a probe microphone inside the ear canal instead of a microphone outside the ear canal entrance.
4.3. Test system

4.3.1. General

Using the results obtained in sections 4.1 and 4.2 a real time speech enhancement system was built in the test car. The system consisted of one speech microphone, two loudspeakers (rear standard speakers), standard car amplifier and a Digital Signal Processor (DSP) system. The microphone was placed in the roof of the car in front of the driver.

To reduce the influence of the background noise, a highpass filter was used. To get the right directionality of the sound a delay was introduced. The microphone signal should be amplified and sent out in the rear speakers, but the driver should not hear him/her self through the loudspeakers.

There were three design parameters in the system:

- Cutoff frequency of the highpass filter.
- Delay time.
- Amplification of the microphone signal.

The system was implemented on a DSP-system. A sketch of the system is shown in Figure 4-12.

![Figure 4-12. Schematic sketch of the test system.](image)

4.3.2. Result test system

The system was tuned to work when the background noise was relatively high (>70 km/h). The microphone amplification should in a real system be varied depending on the background noise level. This parameter was however beyond the scope of this work.
4.3.3. **Conclusions**

The test system showed that it was possible to amplify the signal from the driver position without losing the directionality of the speech. A further development of the system should be varying amplification of the microphone signal depending on background noise level. It was concluded that amplification of the driver’s speech was not necessary when the background noise level was low.

4.4. **Hearing test**

To evaluate if the system (see section 4.3) improved the communication inside the test car an audibility test was performed. For details of hearing tests, see section 2.4.

4.4.1. **Audibility test**

The audibility test was performed in a driving car for three driving conditions; 50 km/h, 90 km/h and 110 km/h. Ten different words were presented for 5 test persons, with the system on and off respectively. The test was a so-called open test, (i.e. the test person knew which words that were to be presented).

The test persons had to judge the audibility on a four-grade scale:

- Did not catch the word (1)
- Weak (2)
- Sufficient (3)
- Good (4)

The test was performed with a loudspeaker that simulated the co-driver. The microphone was placed in the roof in front of the loudspeaker. Otherwise it was the same system as described in section 4.3. The output level from the simulated co-driver, i.e. the loudspeaker was adjusted to normal conversation level. The test persons were given a test sheet (see Appendix B).
4.4.2. Results

In Table 4-2 the result of the audibility test are presented. The result shows that when there is low background noise, i.e. 50 km/h there is no need for a speech enhancement system. When the background noise level raises the need for a speech enhancement system increase. If the audibility test had been performed with music in the background the differences between ON and OFF should probably be higher.

Table 4-2. Results of audibility test.

<table>
<thead>
<tr>
<th>Speed (km/h)</th>
<th>ON</th>
<th>OFF</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>4</td>
<td>3.5</td>
</tr>
<tr>
<td>90</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>110</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

4.4.3. Conclusions

It was shown that the amplification of the microphone signal did not have to be the same all the time. In driving conditions when the background noise level is low, no or small amplification is needed.
Chapter 5 Summary of results

5. SUMMARY OF RESULTS

Section 4.1
The background noise inside a driving car is dominated by low frequency noise. Three methods were tested to reduce the background noise in a microphone signal; (1) microphone array, (2) LMS-algorithm, (3) highpass filter.

1. The microphone arrays that were tested did not manage to reduce noise in the signals efficiently.

2. The LMS-algorithm was never evaluated due to the poor coherence between the reference microphone and the monitor microphone.

3. A highpass filter with cut-off frequency 300 Hz suppressed the background noise in the signal without affecting the speech negatively.

Section 4.2
To get the right directionality of the sound, the law of the first wave front and filtering of the microphone signal was tested and evaluated. A 10 ms delay on the microphone signal sent out in the rear car loudspeaker was enough for the test persons to believe that all sound came from the driver seat.

Section 4.3
From the results in section 4.1 and 4.2 a DSP-system was implemented, in the test car. The system was tuned to work when the background noise was relatively high (>90 km/h).

Section 4.4
The audibility test showed that the speech enhancement system worked for all three driving conditions studied, i.e. 50 km/h, 90 km/h and 110 km/h. The test also showed that it was not necessary to use the system when it was low background noise levels in the car, i.e. 50 km/h.
6. CONCLUSIONS

The use of the car speech enhancement system developed in this thesis work has shown to have potential of increasing the audibility in the passenger compartment. The implemented system worked for several driving condition when the background noise level was high.

The audibility was increased from 3 to 3.6 (20 \%) on a four-grade scale when the car was driving on a motorway, i.e. 110 km/h and in absent of music.
7. RECOMMENDATIONS

Further development and evaluation of the speech enhancement system would include:

- Implementation of a system that follows the background noise level.
- Looking at the possibility of amplifying some frequencies more than others.
- Examination of the frequency response of the loudspeaker in the test car.
- Implementation of the speech enhancement system also for the co-driver.
8. REFERENCES


APPENDIX A Test Linear microphone array

Figure A-1. Loudspeaker positions in test 1

Figure A-2. Loudspeaker positions in test 2
Figure A-3. Linear microphone array with d=7 cm.

Figure A-4. Two-dimensional microphone array with d=7 cm.
Figure A-5. Linear microphone array with d=25 cm.
APPENDIX B Test sheet for the hearing test

Table B-1. Test sheet handed out to the test persons.

<table>
<thead>
<tr>
<th>Word</th>
<th>Test person:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>50 km/h</td>
</tr>
<tr>
<td></td>
<td>ON</td>
</tr>
<tr>
<td>Vantar</td>
<td></td>
</tr>
<tr>
<td>Ägde</td>
<td></td>
</tr>
<tr>
<td>Nio</td>
<td></td>
</tr>
<tr>
<td>Svante</td>
<td></td>
</tr>
<tr>
<td>Tog</td>
<td></td>
</tr>
<tr>
<td>Elsa</td>
<td></td>
</tr>
<tr>
<td>Ringar</td>
<td></td>
</tr>
<tr>
<td>Fina</td>
<td></td>
</tr>
<tr>
<td>Ormar</td>
<td></td>
</tr>
<tr>
<td>Tolv</td>
<td></td>
</tr>
</tbody>
</table>

**Test Instructions:**
You will be presented 10 words for three different driving conditions. Each time you hear a word you shall examine the audibility of the word according to the following instructions: Mark the word with:

1 if you did not catch the word
2 if you had trouble to catch the word
3 if you catch the word
4 if it was easy to catch the word
På svenska

Detta dokument hålls tillgängligt på Internet – eller dess framtida ersättare – under en längre tid från publiceringsdatum under förutsättning att inga extra-ordinära omständigheter uppstår.

Tillgång till dokumentet innebär tillstånd för var och en att läsa, ladda ner, skriva ut enstaka kopior för enskilt bruk och att använda det oförändrat för ickekommersiell forskning och för undervisning. Överföring av upphovsrätten vid en senare tidpunkt kan inte upphäva detta tillstånd. All annan användning av dokumentet kräver upphovsmannens medgivande. För att garantera äktheten, säkerheten och tillgängligheten finns det lösningar av teknisk och administrativ art.

Upphovsmannens ideella rätt innefattar rätt att bli nämnt som upphovsman i den omfattning som god sed kräver vid användning av dokumentet på ovan beskrivna sätt samt skydd mot att dokumentet ändras eller presenteras i sådan form eller i sådant sammanhang som är kränkande för upphovsmannens litterära eller konstnärliga anseende eller egenart.

För ytterligare information om Linköping University Electronic Press se fölagets hemsida http://www.ep.liu.se/

In English

The publishers will keep this document online on the Internet - or its possible replacement - for a considerable time from the date of publication barring exceptional circumstances.

The online availability of the document implies a permanent permission for anyone to read, to download, to print out single copies for your own use and to use it unchanged for any non-commercial research and educational purpose. Subsequent transfers of copyright cannot revoke this permission. All other uses of the document are conditional on the consent of the copyright owner. The publisher has taken technical and administrative measures to assure authenticity, security and accessibility.

According to intellectual property law the author has the right to be mentioned when his/her work is accessed as described above and to be protected against infringement.

For additional information about the Linköping University Electronic Press and its procedures for publication and for assurance of document integrity, please refer to its WWW home page: http://www.ep.liu.se/

© [Mårten Nygren]