Discrimination of short frequency glides depending on reverberation

Bachelor's Thesis

Benedikt Brands

Supervisor: DI Marian Weger
Assessor: O.Univ.-Prof. Mag.art. DI Dr.techn. Robert Höldrich

Graz, April 16, 2019

institut für elektronische musik und akustik

Abstract

Differences in phase or group-delay are crucial properties that help the ear to discriminate between short sounds. To obtain further insights in the ability of short-time discrimination of the auditory system, a forced-choice adaptive listening experiment was performed. The subjects were exposed to chirp-like sounds varying in length, direction of the chirp and added reverberation. These chirp-like sounds were rendered by filtering test-signals with an all-pass filter with logarithmic group-delay. The group-delay was varied and the just notable difference (JND) between the logarithmic group-delay of two stimuli was measured. The results depict that by adding more reverberation, the discrimination decreased. Furthermore, upward chirp-like sounds were discriminable in a better way than downward chirp-like sounds and so were short versus long stimulus lengths. The influence of direction was only significant for short durations.

Eine wichtige Eigenschaft, die das Gehör zur Unterscheidung kurzer Schalle nutzt sind Phasen- und Gruppenlaufzeitunterschiede. Um zu weiteren Einblicken in die Mechanismen der Unterscheidbarkeit zu gelangen, wurde ein adaptiver forced-choice Hörtest durchgeführt. Den Testpersonen wurden chirp-ähnliche Stimuli vorgespielt, welche in Länge, Richtung des Frequenz Sweeps und hinzugefügtem Nachhall variierten. Die Stimuli wurden durch Filtern von Testsignalen mit einem Allpass Filter mit logarithmischer Gruppenlaufzeit realisiert. Diese Gruppenlaufzeit wurde adaptiv variiert und somit die Unterscheidbarkeitsschwelle in logarithmischer Gruppenlaufzeit zwischen zwei Stimuli gemessen. Die Ergebnisse zeigen, dass das Hinzufügen von Nachhall die Unterscheidbarkeit verschlechtert, dass aufsteigende Sweeps besser als absteigende unterscheidbar sind, und dass kurze Stimuli ebenfalls besser als lange Stimuli unterscheidbar sin. Weiterhin war der Einfluss der Richtung nur für kurze Signaldauer signifikant.

Contents

1 Introduction

To detect differences in signals the auditory system makes usage of different mechanisms corresponding to different signal properties. While a change in the magnitude spectrum produces an obvious change in timbre, differences in the phase or group-delay spectrum create a more subtle change. Especially at signal lengths of a few milliseconds where the ear's spectrum analysis mechanism gets inaccurate, changes in phase and group-delay are of great importance for discrimination.

In our everyday life, sound reflections are present immanently. We therefore find ourselves in reverberant environments. This reverberation is crucial for our spatial perception but also has detrimental effects on some mechanisms of auditory perception. In general it has an important role in many psychoacoustical theories.

With those two topics in mind, we focus on signals of short duration (only a few milliseconds) which exhibit the same energy spectrum. Discrimination between signals within this class can therefore only be based on differences in duration, phase and group-delay spectrum, or other complex mechanisms of the auditory system. Adding to that we examine the effect, reverberation may have on that discrimination.

2 Literature Survey

There already has been some research on discrimination of noise and tone bursts [1], or discrimination of nonlinear frequency glides [12]. Also studies to investigate the importance of phase in auditory theory and some studies about the role of reverberation in the process of recognizing sounds were made.
Listening experiments showed that changes in the relative phase among the tones of a complex waveform (under steady-state conditions) produced discriminable differences in its quality [2].
Patterson and Green used so-called Huffman sequences, a technique to generate transient signals with identical energy spectra but different phase spectra [11]. The results suggested that the ear can discriminate differences in temporal order as small as 2.5ms. The generation of the signals was accomplished by filtering an impulse with an all-pass filter which has a constant phase characteristic except for some frequency bands where it changes rapidly by 2π rad due to singularities of the filter. It was observed that for different frequency spans between the singularities the discriminability at low frequencies was essentially perfect while for higher frequencies it decreased with lower spans. They interpreted this as a consequence of the concept of critical bands (CB).

2.1 Auditory filters and the travelling wave theory

The auditory system divides the processing of the perceptible frequencies into a number of frequency bands, the so called Critical Bands (CBs). Many studies focused on psychoacoustic-based spectral measures to obtain knowledge about the nature of the auditory filters which specify these bands. One way was to use narrowband maskers and probe tones to examine the filters [16]. This data gathered in measurements was used to establish the so called Bark scale which divides the perceptible frequency spectrum into 24 CBs [15]. This theory was later revised by Moore and Glasberg who introduced the concept of the Equivalent Rectangular Bandwidth (ERB) which was measured using the notch-noise-method [10], [8].
Both theories have in common that above 500Hz, the bandwidths of the auditory filter increase with frequency in a logarithmic fashion. The theory of the division into CBs or ERBs can be explained by the frequency-to-place transformation in the cochlear mechanics (tonotopy) [13]. It states, that each frequency component excites a specific region on the basilar membrane corresponding to a specific auditory filter (CB or ERB). For high

frequencies, the neural response on the basilar membrane has its maximum near the oval window. For low frequencies, this maximum is located near the apex of the cochlea. Figure 1 illustrates the relations between the Bark scale, the place on the basilar membrane, the pitch mel scale and a logarithmic frequency scale.

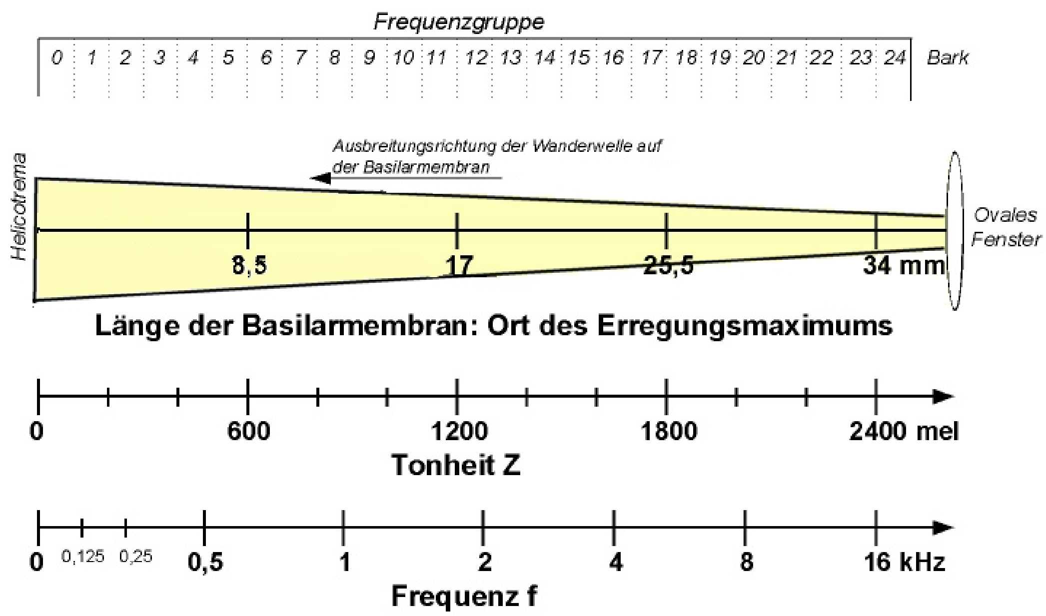

Figure 1 – From top to bottom: Relations between (1) the Bark scale, (2) the place on the basilar membrane, (3) the pitch mel scale and (4) a logarithmic frequency scale, Source: https://de.wikipedia.org/wiki/Bark-Skala

In cochlear mechanics, the concept of the travelling wave on which Von Békésy carried out intense studies is of great importance [14]. It states, that a signal entering the cochlea spreads out as a displacement wave from the base to the apex of the cochlea. Considering the place-theory this yields a delayed processing of low frequencies since regions near the oval window are excited first and regions more far away later. Hence the signal processing inside the cochlea can be regarded as a complex filter bank with an overall, presumably nonlinear group-delay.

2.2 The effect of reverberation on identification and discrimination between sounds

As summarized above, the mechanics of hearing – especially those responsible for discriminability – are well described in terms of frequency- and phase-domain processing. The role of reverberation in those mechanisms is of a more complex nature.
In general, performance deteriorates when adding reverberation but some features of it may also improve acoustical or psycho-acoustical quality measures. For instance, late reflections impair the intelligibility of speech whereas early reflections (typically those reflections which impact until 50-80ms after the direct sound) enhance the effective signal-to-noise-ratio (SNR) and thus enhance the intelligibility of speech [7].
Following Dombois and Eckel another constructive feature is that reverberation may facilitate discrimination between short transient sounds due to the prolongation of these sounds in a neutral way [3, p. 315].
Koumura et. al suggested that the ear is capable to adapt to a reverberant environment and to compensate the effect of reverberation. For example, the impairment of the perception of a word due to reverberation is reduced when presenting it in context of a whole carrier sentence. Furthermore it was observed that participants would adapt to constant reverberation in a listening experiment where identification of the material of an object was examined by listening to the sound generated by impact of the object [5]. The same authors observed in another study that the adaptation to reverberation in a speech scenario does not improve the perception of impact sounds under the same reverberant conditions. This implies that for the same reverberation scenario, the adaptation must be learnt separately for different types of stimuli [6].

A thought experiment may illustrate an interesting approach for the role of reverberation in the ability to identify short sounds:

> *In an acoustically dead room, a listener may not be able name the origin of a short every-day sound such as the clicking of two billiard balls. Does the recognizability improve when reverberation is added?*

This experiment suggests that short every-day sounds are perceived as click-like sounds which may be discriminable but not recognizable and that the effect of blurring over time accompanied by reverberation is crucial fot identifying it. However there are no studies on both effects and the current one examines the effect of reverberation on discriminability.

3 Listening Experiment

Based on the observations made in the discussed literature, a listening experiment was established to obtain further knowledge on the ability to discriminate between short signals with different group-delay spectra and the effect of reverberation on it.

Involving the experiments of Patterson and Green, short signals with identical magnitude spectra are used. Unlike their test signal implementation, for the signals used in this experiment the differences were not set in the phase but in the group-delay spectrum. Using a continuous, monotonically increasing or decreasing group-delay results in a time signal which has the form of an up- or downward frequency glide. To obtain a perceived linear pitch glide, the implementation of this signal must be consistent with the theory of critical bands (CB), respectively the theory of equivalent rectangle bandwidths (ERB), so it is proposed to distribute the group-delay in a logarithmic fashion. By allowing for this, it can be assumed that the frequency glide is exciting each CB in an equal time span, so the perceived pitch glide is linear. The magnitude spectrum of this glide is white, so every frequency band has equal power. Thus, the amplitude of the time signal of the frequency glide increases for higher frequencies. The implementation is accomplished by filtering a test signal by an all-pass filter with logarithmic group-delay spectrum $\tau(f)$. Therefore an impulse-like test signal processed by the all-pass filter yields an exponential frequency glide.

To measure the ability of the auditory system to discriminate between short sounds an adaptive listening experiment was carried out. The test signals used were either a Dirac-impulse or a noise burst of base durations T which were filtered by an all-pass filter with logarithmic group-delay spectrum $\tau(f)$ resulting in either a click-like sound (for a Dirac-impulse input) or chirp-like sound (for a noise burst input). The convolution of the test signal with this filter extended the test signal up to the length of $T + \Delta\tau_{AP}$ where $\Delta\tau_{AP}$ is the group-delay span of the filter. The test signals were convolved with the impulse response (IR) of the all-pass filter or the time-reversed IR resulting in an up- or downward chirp-like sound. The last variation was to add one out of two types of reverberation or none yielding three reverberation scenarios. The independent parameter which was adaptively varied was the length of the filter $\Delta\tau_{AP}$ resulting in a variation of the length $T + \Delta\tau_{AP}$ of the chirp-like sound. Since all sounds were filtered, the just notable difference (JND) is the difference of the target group-delay and the reference group-delay $\Delta\tau = \Delta\tau_{target} - \Delta\tau_0$.

In summary there were three independent variables that were tested for their influence on the JND in time $\Delta\tau$: the base length of the test signal T, the spectral direction of the group-delay and the reverberation scenario. Therefore the JND that was measured can be regarded having the dimension "difference in logarithmic group-delay" due to the nature of the filter.

3.1 Generation of Test signal

In attempt to get knowledge of the discriminability two base durations T were selected to represent two short-time signal extremes: a very short one to measure the threshold for the discriminability of *click-like* sounds and a mid-short one to measure the threshold for the discriminability of short *chirp-like* sounds, all of which have the same magnitude spectra.
As Moore summarized, the smallest detectable increment in duration ΔT increases for base durations T exceeding 10 ms [9, p. 171]. Therefore a base duration of T_{long} = 10 ms was chosen for the *chirp-like* sounds which was filled with uniformly distributed noise.
To realize the very short *click-like* sound, the other length was set to one sample to obtain an ideal Dirac impulse.

3.2 Implementation of the all-pass filter with logarithmic group-delay spectrum

The all-pass filter was generated in Matlab. The procedure started with the design of the magnitude spectrum $|H_{bandpass}(f)|$ being equal to one over the predefined bandpass spectrum (f_{start}=500Hz, f_{stop} =16kHz) and zero everywhere else. Then a Hann-window which had the length of the wanted signal but at least the length of a prespecified minimum of 3ms was applied in the frequency-domain via circular convolution. In a next step, the bandstop regions were set to zero again. The result is a bandpass filter. A frequency-dependent phase $\phi(f)$ was added to this filter, leading to the desired (band-limited) all-pass filter $H_{chirp}(f)$.

The group-delay spectrum $\tau(f)$ with a delay span over the bandwidth $f_{start} - f_{stop}$ of $\Delta\tau_{AP}$ was created using

$$\tau(f) = \Delta\tau_{AP} \cdot \frac{\log f - \log f_{start}}{\log f_{stop} - \log f_{start}} \tag{1}$$

Since the group-delay is defined as the negative derivative of the (unwrapped) phase spectrum

$$\tau(\omega) = -\frac{d\phi(\omega)}{d\omega} \tag{2}$$

we obtain, using $\omega = 2\pi f$, the phase spectrum:

$$\phi(f) = -2\pi \int \tau(f) df \tag{3}$$

The transfer function of the wanted filter was now rendered as follows:

$$H_{chirp}(f) = |H_{bandpass}(f)| \cdot e^{j\phi(f)} \tag{4}$$

Via inverse FFT, the impulse response of the filter $\hat{c}(t)$ was obtained

$$\hat{c}(t) = \text{IFFT}\{H_{chirp}(f)\} \tag{5}$$

Then a Tukey-window with a fade time of 10ms and a taper time of 1.5 times the duration of the signal was applied on the IR $\hat{c}(t)$. For short signals, the energy is important for the perceived loudness. Since an equal loudness for each stimulus is desired, eventually an energy normalization of the IR was applied and the final IR $c(t)$ was rendered:

$$c(t) = \frac{\hat{c}(t)}{\sqrt{\int \hat{c}(t)^2 dt}} \tag{6}$$

The generated group-delay spectrum and phase spectrum using equation (1) and (2) are shown in Fig. 2. The magnitude and group-delay spectra of the normalized all-pass filter after applying equation (4) and (6) are shown in Fig. 3. The normalized impulse response $c(t)$ of the filter can be seen in Fig. 4. Note that due to the trade-off between an ideal rectangular bandwidth magnitude spectrum and the realization with a finite impulse response (FIR) filter, the desired magnitude and group-delay properties can be fulfilled within the specified bandwidth but deviate above and below it. Since there is a huge drop in magnitude outside the specified bandwidth, the corresponding anomalies in the group-delay spectrum can be disregarded.

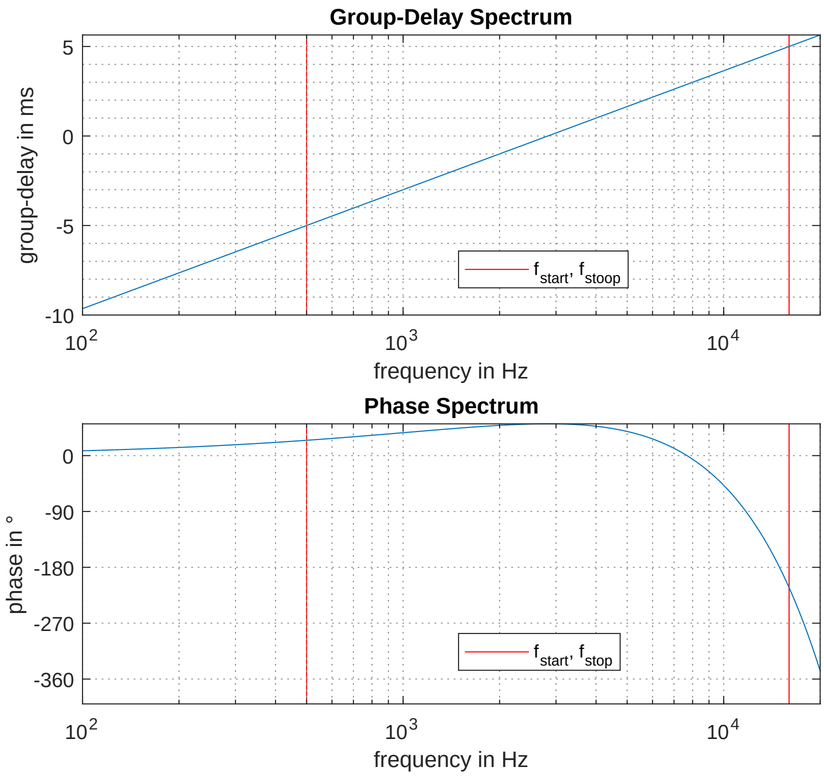

Figure 2 – Analytically defined group-delay and phase spectrum for $\Delta\tau_{AP}$ =10ms over a bandwidth of $f_{start} = 500$Hz to $f_{stop} = 16$kHz

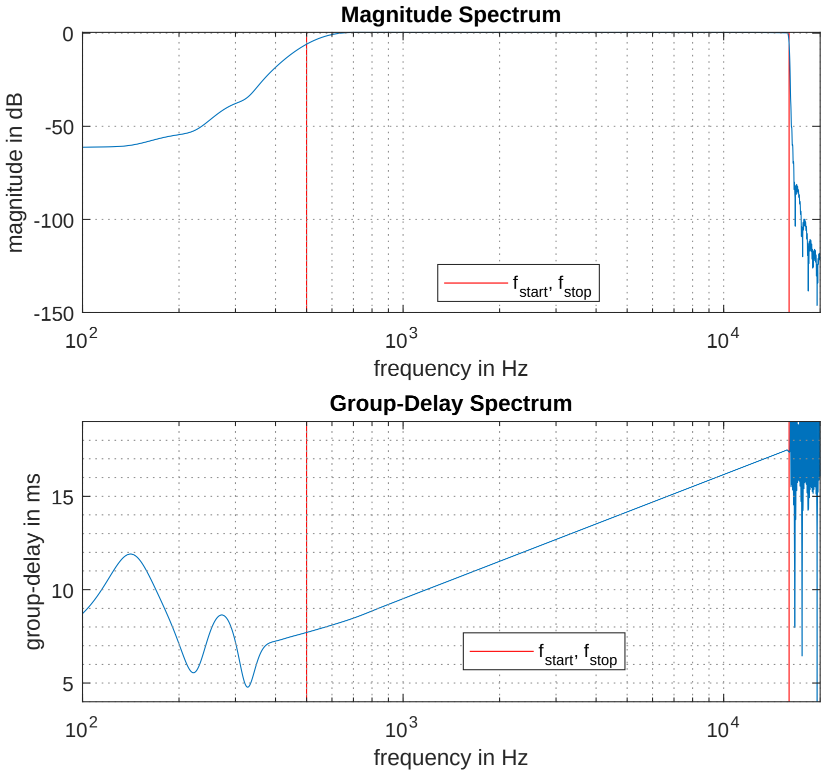

Figure 3 – Magnitude and group-delay spectrum of the normalized all-pass filter with logarithmic group-delay span of $\Delta\tau_{AP}$ =10ms over a bandwidth of f_{start} = 500Hz to f_{stop} = 16kHz

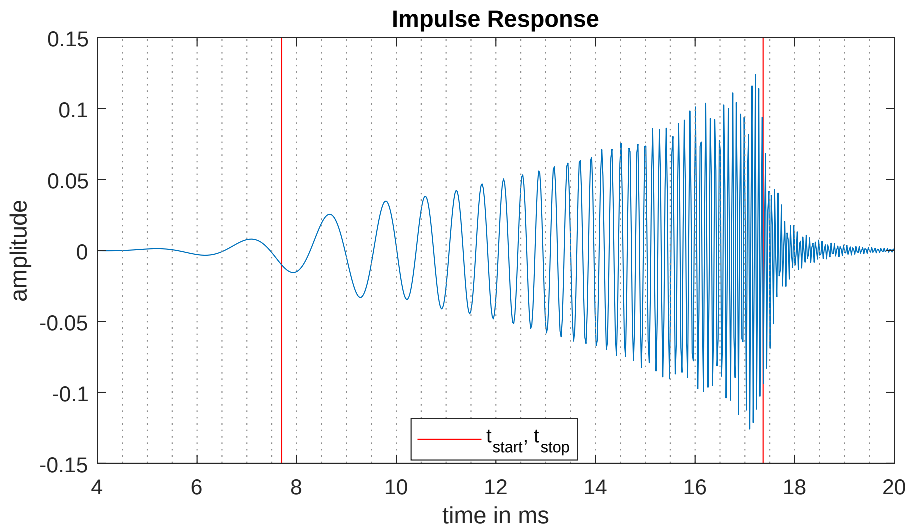

Figure 4 – Impulse response of the all-pass filter with logarithmic group-delay spectrum for $\Delta\tau_{AP}$ =10ms. t_{start} and t_{stop} indicate the time at which the momentary frequency $f(t)$ equals f_{start} = 500Hz and f_{stop} = 16kHz, respectively

3.3 Reverberation scenarios

Three reverberation scenarios were defined: The first one having no reverberation (*dry*), the second one having a low reverberation as in a well designed conference room (*room*) and the third one as a multi-purpose acoustics laboratory and small concert hall having medium long reverberation (*hall*). For the first scenario the dry chirp was used without further modification. For the second scenario a measurement of the binaural impulse response of the lecture room Petersgasse using a cube loudspeaker was applied. A picture of the lecture room is shown in Fig 5. In the third scenario a measurement of the impulse response of the IEM CUBE using a icosahedron loudspeaker was used. A picture of the room is shown in Fig. 6. Both measurements were recorded with a Neumann KU100 dummy head microphone. The lecture room had an area of 50m^2 and a volume of approximately 180m^3 and the IEM CUBE an area of 120m^2 and a volume of around 420m^3. The rooms had an estimate reverberation time T_{60} of 0.5s and 0.7s respectively. The reverberation time was obtained by estimating the duration of a 30dB descent in the IR and then extrapolating it up to 60dB.

Figure 5 – Picture of the lecture room Petersgasse [1]

1. copyright Nils Meyer-Kahlen

Figure 6 – Picture of the IEM CUBE [2]

3.4 Stimuli, method and apparatus

As default values for the test signal and the all-pass filter the following values were set:

— test signal base durations: $T_{short} = 1$ sample @ 44.1kHz and $T_{long} = 10$ms

— spectrum of frequency glide: $f_{start} = 500$Hz, $f_{stop} = 16$kHz

— reference group-delay span: $\Delta\tau_0 = 10$ms

— initial group-delay span of target: $\Delta\tau_{target} = 20$ms

Thresholds were measured using a forced-choice adaptive triangle test. Three stimuli were played with a break time which completed the length of one stimulus to 700ms each, so that the perceived rhythm pattern was constant over all trials. There were three independent variables: duration with two levels (*short, long*), direction with two levels (*up, down*), and reverberation type with three levels (*dry, room* and *hall*). This leads to 2x2x3=12 conditions. These twelve conditions were presented to the participants in randomized order. For the triangle test two stimuli had the reference length $T + \Delta\tau_0$ and the target was of length $T + \Delta\tau_{target}$. The playback order of these three stimuli was randomized for each trial and the subject had to detect the target stimulus with help of a replay button. The graphic user interface for the listening experiment is shown in Fig. 7.

2. copyright Kunstuniversität Graz https://www.kug.ac.at/

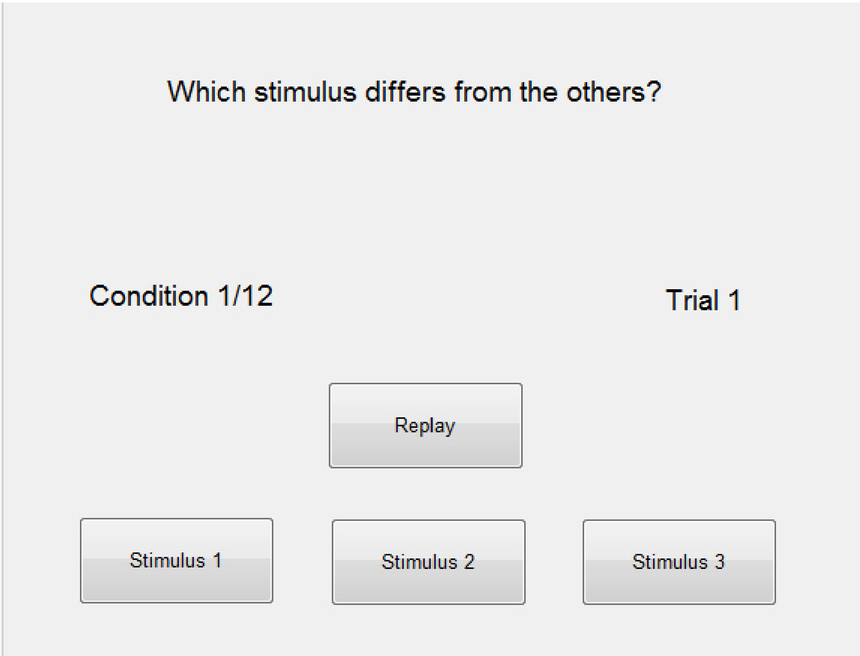

Figure 7 – Graphic user interface (GUI) for the listening experiment with question, condition and trial counter, replay button and three answering options. For the correct answer the GUI shortly blinked in green, for the wrong answer it blinked in red.

It also gave feedback to the participant's answer with green blinking for the correct and red for the wrong answer. The difference in duration $\Delta\tau = \Delta\tau_{target} - \Delta\tau_0$ was initialised with 10ms and was varied adaptively following a three-up one-down rule with a fixed stepsize (50, 50, 50, 25, 25, 10, 10, 10, 8, 8, 5, 5 percent of $\Delta\tau$). A measurement ended if either 100 trials were passed or 11 reversals were done. For calculating the median value of the difference in duration the first five reversals were ignored, hence the median value over the last six reversals was taken as a single threshold. An exemplary plot of the progression of $\Delta\tau$ over the trial number is shown in Figure 8.

The stimuli were generated by a Matlab function. The listening experiment was executed and the subject responses recorded by a Matlab script based on a script by Arthur Lugtigheid[3]. The impulse response of the IEM CUBE was provided by Univ. Prof. Dr. phil. Gerhard Eckel and the impulse response of the Hörsaal Petersgasse was provided by Nils Meyer-Kahlen. The digital stimuli were converted into analog signals by means of a stereo 24-bit D/A converter (Lexicon Alpha) at a sampling rate of 44.1kHz. The stimuli were then presented to the listener over headphones (Audio-Technica ATH-M50).

3. https://github.com/lugtigheid/matlabstaircase

12 subjects aged between 21 and 33 years participated in the experiment. All participants had normal hearing. Eleven were sound engineers (10 students and one university staff member), and one was a musician. The subjects obtained no explicit training for the experiment but all of them can be regarded as sound professionals.

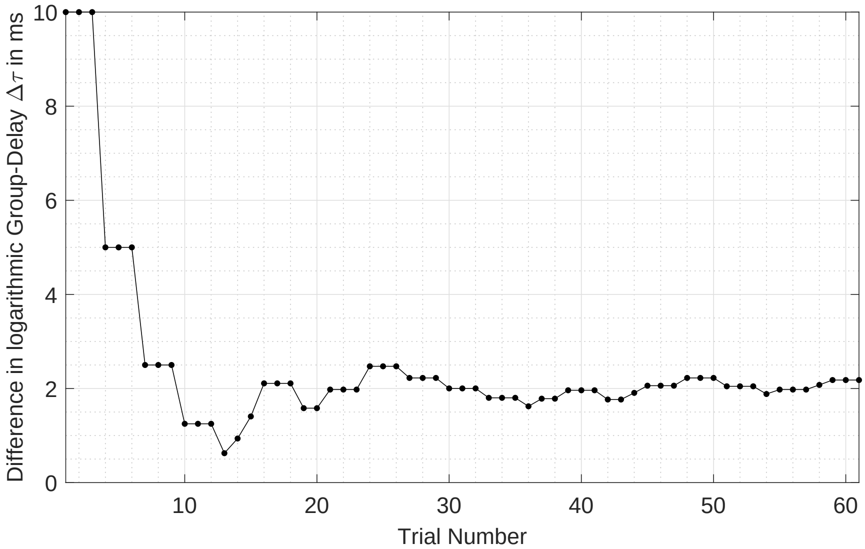

Figure 8 – Exemplary plot of the difference in logarithmic group-delay $\Delta\tau$ over the trial number for participant #7

3.5 Results

For every participant a threshold of $\Delta\tau$ for each of the 12 conditions was recorded. Before calculating mean-values, the raw data was tested for outliers by means of Grubb's test [4]. Three of the 12 conditions contained one outlier (out of 12 participants). These three outliers were replaced by the median of the other 11 participants, respectively. The average JND of $\Delta\tau$ for each condition is visualized in Table 1, respectively Fig. 9 with error bars indicating 95% confidence intervals.

The Table and the Figure indicate, that

- there is an overall trend that for increasing reverberation, the JNDs ascend
- the minimum JND $\Delta\tau$=2.7ms occurs at the condition *dry/short/up* and the maximum JND $\Delta\tau$=6.5ms at the condition *hall/short/down* and *hall/long/down*
- reverberation scenario *dry* exhibits a smaller JND than *room* which exhibits a smaller JND than *hall* except for the condition *long/down*, where *dry* has a greater JND than *room*
- duration *short* exhibits a smaller JND than *long* except for the condition *room/down* where *short* has a greater JND than *long*
- direction *up* exhibits a smaller JND than *down* except for the condition *hall/long* where the JNDs are equal

	JNDs in ms			
	short		long	
	up	down	up	down
dry	2.7	3.5	3.9	5.3
room	3.8	5.2	4.9	5.1
hall	5.1	6.1	6.5	6.5

Table 1 – averaged JNDs for the three reverberation conditions, the two base length conditions and the two direction conditions for a reference group-delay span of $\Delta\tau_0 = 10$ms over the bandwidth f_{start} =500Hz, f_{stop} =16kHz

3.6 Statistical analysis of the results

We proceed with a statistical analysis of the results. Firstly the data is tested for normality to assess if the analysis of variance (ANOVA) is applicable. If it is, we examine which main factors and interactions are significant. Eventually pairwise t-tests of the conditions are applied.

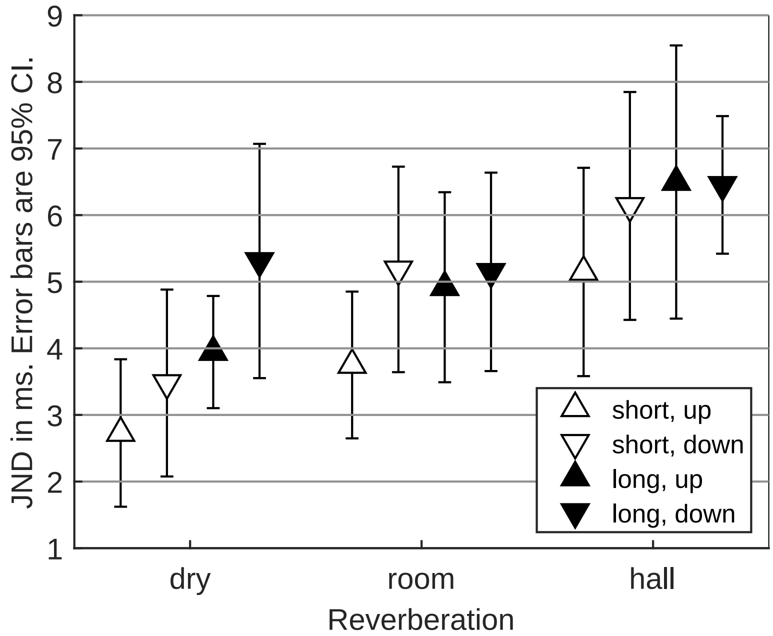

Figure 9 – JNDs of $\Delta\tau$ and 95% confidence interval (CI) for the three reverberation conditions, the two duration conditions and the two direction conditions for a reference group-delay span of $\Delta\tau_0 = 10$ms over the bandwidth f_{start} =500Hz, f_{stop} =16kHz

Test for normally distributed data

Null hypothesis: *"the data is normally distributed"*

The data is tested for normality by means of a Matlab implementation of the Lilliefors-Test. The test returns a p-value which helps to decide if normally-distributed data is available. At the 5% significance level, $p \geq 0.05$ indicates that the null hypothesis cannot be rejected and $p < 0.05$ indicates, that the null hypothesis can be rejected, so the data is non-normal. For the results of the listening experiment, only one condition (duration *long*, direction *up*, reverberation scenario *room*) is non-normal ($p = 0.0222$). In total $\frac{1}{12} \approx 8.33\%$ of all conditions is non-normal, so we decide to apply the ANOVA.

ANOVA

Null hypothesis: *"the main variable has no influence on the JND"*

Using the analysis of variance with three way repeated measures, the influences of the different variables on the JNDs were examined. As a significance level for rejecting the null hypothesis 5% was chosen. The results are shown in Table 2. The analysis stated that for all main effects the null hypothesis can be rejected, so all have main effects have a significant influence on the JND:

— duration: $F(1,11) = 22.630$, $p = 0.001$
— direction: $F(1,11) = 18.543$, $p = 0.001$
— reverberation: $F(2,22) = 49.655$, $p < 0.001$

Null hypothesis: *"the interactions of the variables have no influence on the JND"*

An examination on the interaction of the variables showed that none of the possible interactions between the variables influenced the JND significantly (see Table 2: all p-values are greater than 5%, so the null hypothesis cannot be rejected).

Pairwise t-test

Null hypothesis: *"the given pair of variables exhibit no different JNDs"*

The pairwise t-test of the averaged data showed that:

— reverberation scenario *room* (mean: 4.75ms) led to a significantly higher JND ($t(11) = -4.087$, $p = 0.002$) than *dry* (mean: 3.87ms)
— reverberation scenario *hall* (mean: 6.06ms) led to a significantly higher JND ($t(11) = -6.807$, $p < 0.001$) than *room*
— duration *long* (mean 5.38ms) led to a significantly higher JND ($t(71) = -4.563, p < 0.001$) than duration *short* (mean 4.40 ms)
— direction *down* (mean 5.29ms) led to a significantly higher JND ($t(71) = -4.563, p < 0.001$) than *up* (mean 4.50ms)

For the other averaged results, the null hypothesis could not be rejected.

A pooled t-test over the reverberation conditions showed that:

— for the direction *up*, the JNDs for duration *short* differed significantly ($t(35) = -6.038, p < 001$) to those for *long*
— for the duration *short*, the JNDs for direction *up* differed significantly ($t(35) = -5.164, p < 001$) to those for *down*
— for the direction *down*, the JNDs for duration *short* differed significantly ($t(35) = -2.234, p = 0.032$) to those for *long*.
It must be noted that this p-value is relatively close to the 5% significance level and might not withstand a more critical statistical methods
— for the duration *long*, the JNDs for direction *up* did not differ significantly ($t(35) = -1.893, p = 0.067$) to those for *down*

	SumSq	DF	MeanSq	F	p-value
duration	3.415e-05	1	3.415e-05	22.63	<0.001
direction	2.2378e-05	1	2.2378e-05	18.543	0.001
reverberation	0.00011682	2	5.841e-05	49.655	<0.001
duration x direction	2.6295e-06	1	2.6295e-06	3.1609	0.103
duration x reverberation	5.8691e-06	2	2.9346e-06	2.7768	0.084
direction x reverberation	2.0854e-06	2	1.0427e-06	1.5559	0.233
duration x direction x reverberation	6.0717e-06	2	3.0358e-06	2.3557	0.118

Table 2 – Results of three way repeated measures ANOVA

3.7 Discussion

Effect of reverberation

As the results show, reverberation worsens the discrimination of short sounds having the same magnitude spectrum and different logarithmic group-delay spectra since the JNDs deteriorate significantly with longer reverberation. This observations objects the statement of Dombois and Eckel, that reverberation may produce a better discrimination between transient signals.

Effect of duration and direction

The results indicate that the listeners could distinguish short sounds (of length $T_{short} + \Delta\tau_0 + \Delta\tau \approx$10ms+$\Delta\tau$) better than long sounds (of length $T_{long} + \Delta\tau_0 + \Delta\tau =$20ms+$\Delta\tau$) and that the ability to distinguish between upward frequency glides is better than for downward frequency glides. The latter observation supports the theory that the processing time of the inner ear is dependent on frequency. One might draw the false conclusion that the inner ear takes more time to process high frequencies than low frequencies resulting in a stretching in time for upward frequency glides and a compression in time for downward frequency glides. In contradiction, the theory of the travelling wave (see Sec 2.1, page 3) states the exact opposite: for an upward frequency glide the nature of the cochlear frequency processing yields a compression, respectively a stretching in time for a downward, respectively an upward frequency glide (see Fig. 10). Adding to that, the pairwise t-test of the pooled data shows, that the effect of direction on the JNDs is only significant for short base durations.

The latter observation allows the hypothesis, that the ear is able to differentiate by means of differences in timbre (reported as being like "tic" and "tock") in a better way than it is by means of recognizing a difference in two frequency glides. Thus, sounds of short duration which are recognized as click-like sounds produce a better discrimination than sounds of long duration which are already recognized as chirp-like sounds. The threshold between these two mechanisms of discrimination is shifted by the process of distortion in time due to the travelling wave theory. This procedure benefits discrimination for *short up*ward frequency glides since the glide is compressed to a click-like sound (see Fig. 10 (a)) and disadvantages *short down*ward glides since the glide is stretched to a chirp-like sound (see Fig. 10 (b)).

Regarding the *Weber Law* another hypothesis can be established. The *Weber Law* states, that the just notable difference ΔS is proportional to the initial stimuli intensity S, so the so called *Weber Fraction* is constant: $\frac{\Delta S}{S}$ =const. For the JNDs and initial stimulus lengths observed in this experiment, this fraction is not constant. But regarding the group-delay properties of the ear and its consequences - the compression and stretching in time (see Fig. 10) - the resulting perceived stimulus lengths may result in a constant *Weber Fraction*. However, the fact that the *Weber Law* fails in the vicinity of the absolute detection threshold has to be considered, since the used stimulus lengths were very short.

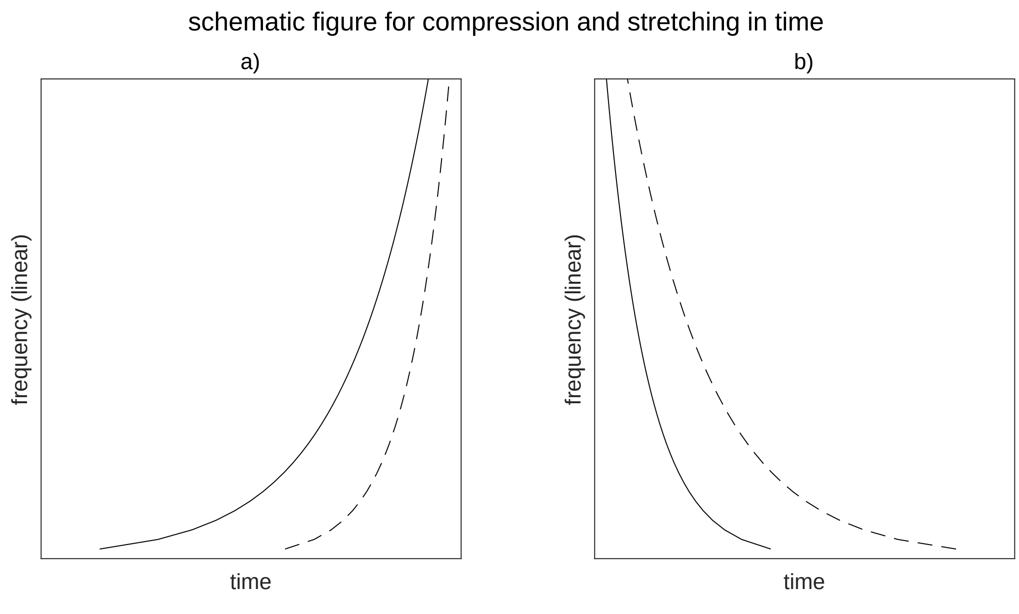

Figure 10 – schematic figure of distortion in time of frequency glides due to delaying frequency processing time in the inner ear: (a) compression in time for an upward frequency glide, (b) stretching in time for a downward frequency glide. Drawn through line: frequency glide before processing, dashed line: frequency glide after processing

4 Conclusion and outlook

The aim of this study was to get further insights in the process of hearing regarding the discrimination of short sounds having identical magnitude spectra and the effect of reverberation on this ability. The sounds used were short click- and chirp-like sounds which were implemented using an all-pass filter with logarithmic group-delay spectrum. There were three fixed conditions (duration T: *short* (base duration of one sample at 44.1kHz) and *long* (base duration of 10ms); direction: *up* (upward frequency glide) and *down* (downward frequency glide); reverberation scenario: *dry*(no reverberation), *room* (little reverberation) and *hall* (medium reverberation)) resulting in 2x2x3=12 conditions. The independent variable was the length and therefore the group-delay span of the filter $\Delta\tau_{AP}$ which elongated the test signal (of length T) to the length $T + \Delta\tau_{AP}$. In a forced choice adaptive triangle test the just noticeable difference (JND) in logarithmic group-delay between the target and the reference group-delay span $\Delta\tau = \Delta\tau_{target} - \Delta\tau_0$ was measured. The reference group-delay was $\Delta\tau_0$ =10ms

The statistical analysis showed that all conditions had a significant effect on discrimination. It furthermore indicated that the effect of reverberation worsens discrimination and that upward frequency glides exhibit a higher JND than downward frequency glides and that the duration *long* used in this experiment led to a higher JND than the duration *short*. Further analysis revealed that only for short signal lengths the JNDs for the condition direction were significantly different.

That latter observation that the ear can discriminate between short upward frequency glides better than between short downward frequency glides and the theory that in the inner ear, high frequencies are processed faster than low frequencies, may allow the theory that the ear's ability to discriminate between click-like sounds is better than that for discriminating between chirp-like sounds.

To entirely understand the mechanism of hearing all of the many stages of the human ear must be studied and conceived. The processing in the cochlea is only one out of many but still provides some procedures to be investigated. As this study highlighted, the time-dependent frequency processing may have influence on short-time discrimination. By testing the ear with sounds of specific length and nature some further observations may be gathered.

As mentioned in the Introduction, the listening experiment used in this study was designed to allow a later application for examining the effect of reverberation on the identification of short every-day sounds. In the process of this study there already had been recordings of these sounds in a room with low reverberation. An acoustically dead room was not available but the removal of the reflections in the recordings delivered accurate results. For the reverberant scenarios the IRs used in this study may be used again. Since the question for identification yields many degrees of freedom a carefully selected set of options which does not influence or disturb the listener's choice should be used. In addition, the effect of learning due to repetition should be prevented. These requirements demand a clever test design. For instance, the all-pass filter with logarithmic group-delay spectrum implemented in this study may be used to alienate the recorded sounds and distract from the process of learning.
The knowledge about the process of hearing that may be gathered in those recommended studies could be used in some problems of sonification.

References

[1] Sharon M. Abel. Duration discrimination of noise and tone bursts. *The Journal of the Acoustical Society of America*, 51(4B):1219–1223, 1972.

[2] James H. Craig and Lloyd A. Jeffress. Effect of phase on the quality of a two-component tone. *The Journal of the Acoustical Society of America*, 34(11):1752–1760, 1962.

[3] Florian Dombois and Gerhard Eckel. *The Sonificaion Handbook*. 2011.

[4] Frank E. Grubbs. Procedure for detecting outlying observations in samples. *Technometrics*, 11:53, 04 1974.

[5] Takuya Koumura and Shigeto Furukawa. Context-dependent effect of reverberation on material perception from impact sound. *Scientific Reports*, 7(16455), 11 2017.

[6] Takuya Koumura and Shigeto Furukawa. Do speech contexts induce constancy of material perception based on impact sound under reverberation? *Acta Acustica united with Acustica*, 104:796–799, 09 2018.

[7] Abigail Anne Kressner, Adam Westermann, and Jörg M. Buchholz. The impact of reverberation on speech intelligibility in cochlear implant recipients. *The Journal of the Acoustical Society of America*, 144(2):1113–1122, 2018.

[8] Brian C. J. Moore and Brian R. Glasberg. Suggested formulae for calculating auditory-filter bandwidths and excitation patterns. *The Journal of the Acoustical Society of America*, 74(3):750–753, 1983.

[9] Brian C.J. Moore. *An Introduction to the Psychology of Hearing*. Academic Press, 2003.

[10] Brian C.J. Moore and Brian R. Glasberg. A revision of zwicker's loudness model. *Acta Acustica united with Acustica*, 82:335–345, 03 1996.

[11] James H. Patterson and David M. Green. Discrimination of transient signals having identical energy spectra. *The Journal of the Acoustical Society of America*, 48(4B):894–905, 1970.

[12] Nick Thyer and Doug Mahar. Discrimination of nonlinear frequency glides. *The Journal of the Acoustical Society of America*, 119(5):2929–2936, 2006.

[13] Hartmut Traunmüller. Analytical expressions for the tonotopic sensory scale. *The Journal of the Acoustical Society of America*, 88(1):97–100, 1990.

[14] Georg von Békésy. *Experiments in Hearing*. McGraw-Hill, 1960.

[15] E. Zwicker. Subdivision of the audible frequency range into critical bands (frequenzgruppen). *The Journal of the Acoustical Society of America*, 33(2):248–248, 1961.

[16] Eberhard Zwicker and Bertram Scharf. A model of loudness summation. *Psychological review*, 72:3–26, 02 1965.