

Diploma Thesis

Investigation of psychoacoustic principles for automatic mixdown algorithms

Philipp Aichinger

Supervisors:
Prof. Dr. Gerhard Eckel
Univ. Ass. DI Dr. Alois Sontacchi

University of Music and Performing Arts Graz, Austria
Institute of Electronic Music and Acoustics (IEM)

Investigation of psychoacoustic principles for automatic mixdown algorithms

Philipp Aichinger

(University for Music and Performing Arts Graz, Austria

Institute of Electronic Music and Acoustics (IEM)

philaich@hotmail.com)

Abstract: The diploma thesis at hand describes the implementation of a psychoacoustic hearing model for modeling spectral loudness of sounds in noise, i.e. mixed instrumental sounds. The model accounts for all known physical and physiological effects between the sound source and the cerebral stage. It is used to visualize the spectral behavior of mixed instruments. It is also used to design a perceptual experiment that investigates the subjects' ability to identify different instruments in a masking environment. Different levels of masking are indicated with the aid of the newly defined measure "loudness quotient" that relates masked to unmasked (original) loudness. Instruments with a quotient of 10 % or smaller or with a residual loudness of 1 sone or less are identified to be critical in mixdowns because they cannot be identified adequately. A linear regression model is defined and proved to model the relation between the quotient of loudness and the identification probability of mixed instruments. Besides that a new approach for predicting binaural unmasking is proposed and tested on empirical data. The results of the study can be used for optimizing existing automatic mixdown algorithms, as groundwork for developing new mixing algorithms or for the revision of ordinary mixing console approaches and user interface concepts.

Diese Arbeit beschreibt die Implementierung eines psychoakustischen Gehörmodells, das die spektrale Lautheit maskierter (gemixter) Instrumente modelliert. Das Modell berücksichtigt alle bekannten physikalischen und physiologischen Effekte zwischen der Schallquelle und der cerebralen Verarbeitung. Es wird dazu verwendet, das spektrale Verhalten gemixter Instrumente zu visualisieren. Weiters wird das Modell herangezogen, um ein Experiment zu designen, bei dem die Fähigkeit der Probanden ein maskiertes Instrument zu erkennen untersucht wird. Die verschiedenen Stufen der Maskierung werden durch die neu definierte Maßzahl des Lautheitsquotienten gekennzeichnet, der maskierte und unmaskierte (originale) Lautheit in Relation setzt. Es wird gezeigt, dass Instrumente mit einem Quotienten von 10 % oder kleiner oder mit einer Restlautheit von 1 sone oder weniger im Mix schlecht oder gar nicht identifiziert werden können. Um die Beziehung von Lautheitsquotient und Identifikationswahrscheinlichkeit zu modellieren, wird ein lineares Regressionsmodell definiert und dessen Gültigkeit gezeigt. Weiters wird ein neuer Ansatz binaurales Demaskieren zu modellieren vorgeschlagen und per Vergleich mit empirischen Daten überprüft. Die Ergebnisse der Studie können dazu verwendet werden, bereits existierende automatische Mixdownalgorithmen weiter zu entwickeln. Weiters können die Ergebnisse als Grundlage für die Entwicklung neuer Mixdownalgorithmen oder für die Überarbeitung herkömmlicher Mischpultkonzepte und Ansätze für Mensch-Maschine Schnittstellen verstanden werden.

Keywords: psychoacoustic, hearing model, auditory masking model, masking threshold, loudness model, partial masking, perceptual model, timbre, temporal masking, binaural unmasking, binaural loudness summation, fundamental note, partial tone, discrimination of instruments, audibility, partial loudness, specific loudness, instantaneous loudness, short-term loudness, long-term loudness, instantaneous partial loudness IPL, short-term partial loudness STPL, automatic mixing, identification threshold

Pledge of Integrity

I hereby certify that the work presented in this thesis is my own, that all work performed by others is appropriately declared and cited, and that no sources other than those listed were used.

Place: _____

Date: _____

Signature: _____

Eidesstattliche Erklärung

Ich versichere ehrenwörtlich, dass ich diese Arbeit selbständig verfasst habe, dass sämtliche Arbeiten von Anderen entsprechend gekennzeichnet und mit Quellenangaben versehen sind, und dass ich keine anderen als die angegebenen Quellen benutzt habe.

Ort: _____

Datum: _____

Unterschrift: _____

Content

1 Introduction	5
2 The hearing model	6
2.1 Calibration	7
2.1.1 Incoming signals	7
2.1.2 Outgoing signals	8
2.2 Outer ear filtering	9
2.2.1 Stereo setup	9
2.2.2 Headphones setup	10
2.3 Middle ear filtering	12
2.4 Filter realization	14
2.5 Time-frequency analysis	14
2.6 Cochlear filtering	18
2.6.1 Details and implementation	20
2.7 Masking threshold	26
2.7.1 Experiment: adjusting the model	26
2.8 Loudness of steady sounds in noise	35
2.8.1 Partial or specific loudness	35
2.8.2 Overall or total loudness	39
2.9 Loudness of timevarying sounds in timevarying noise [2]	42
2.9.1 Calculation of short-term loudness	42
2.9.2 Calculation of long-term loudness	43
2.9.3 Instantaneous partial loudness IPL and short-term partial loudness STPL	43
2.10 Binaural summation of loudness	44
2.10.1 Binaural unmasking for two sound sources	47
2.11 Cross-analysis	50
2.12 Restrictions of the model	53
3 Experiment	53
3.1 Motivation	53
3.1.1 Quotient of loudness for masked instruments	53
3.2 Testsignals	54
3.2.1 The spectral shape of the masker	55
3.3 Experimental design	59
3.4 Report	59
3.4.1 Qualifying	59
3.4.2 Overall results	61
3.4.3 Instrumental threshold determination	67
4 Conclusions and perspectives	82

1 Introduction

In [4] authors proposed an algorithm that lets the user decide for a single important instrument to enhance instead of finding a compromise for all instruments. It works with gain attenuation instead of adapting filter settings. It finds a kind of spectral centroid for each instrument and derives spectral similarities only from this single scalar which now can be refined to a precise psychoacoustic model.

The results of the investigations can also be used to draw consequences on ordinary mixing console approaches. Instead of just attenuate or gain the level of an instrument without any perceptual motivated connections between the faders, it seems nearby to create a tool that can do more. One can think of a faderbank that directly lets the user touch the perceptual importance of an instruments. Beside that this mixing tool can ensure that once a desired LQ is found it does not changes in long term meanings again, even when more instruments are added, because the algorithm temporarily updates the consoles settings. Psychoacoustic visualization can be used to show what frequencies of what instruments compete against each other so that the engineer can find better filter settings more easily and faster.

Along the recording chain, the sound field of one or more instruments is transformed from the mechanic domain into the electric domain by one or more electroacoustic transducers (microphones). If you electronically sum these signals to any kind of mixdown format (e.g. Stereo or 5.1), on the one hand the aim may be to picture natural sound fields as close as possible to reality, on the other hand unreal acoustic scenarios are created consciously or unconsciously. The paper at hand develops an evaluating description for different scenarios, relating to tested perception psychological effects.

It is for instance possible to look at the spectrum of a signal and deduce its masking characteristics, i. e. calculate the hearing thresholds that variable concerning time and frequency. If you mix a signal to a second signal, parts of this newly added signal will be audible, while others will remain inaudible. But also the newly added signal works as a masker and covers or reduces parts of the first signal. These proportions are dependent on enormous fluctuation in time and frequency and are graphically and quantitatively visualized by the use of signal theoretical and statistical tools, such as spectrograms and specific loudness graphs. Also the positioning in the panorama of the mixdown has effects on these masking relations and is investigated shortly in section 2.10.

Section 2 describes the theory and implementation of a hearing model given by [1], [2] and [14]. Section 3 describes the design, conduction and results of a listening experiment that determines the relation between the identification probability of masked instruments and the level of masking. Section 4 subsumes the results and lists possibilities of related further research.

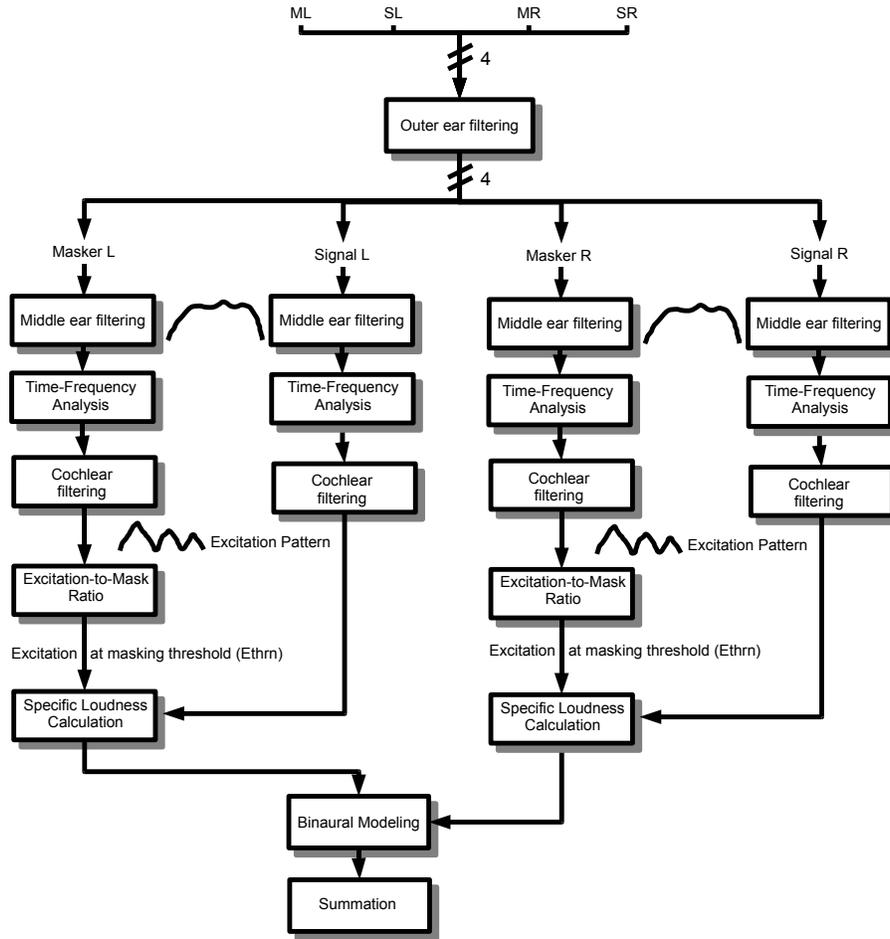


Figure 1: Block diagram of the used hearing model, the inputs SL, ML, SR and MR are signals and maskers of the left and right loudspeaker of the stereo arrangement and will further be explained in section 2.2.1

2 The hearing model

The first step is to find a model that can predict whether a sound embedded in an acoustical environment is audible or not. Further on the model must predict the perceived loudness spanning over the frequency axis (i.e. Specific Loudness N' , [sone/bark]).

Figure 1 shows the structure of the hearing model. Its blocks (i.e. outer ear filtering, middle ear filtering, time-frequency analysis, cochlear or inner ear fil-

tering, excitation-to-mask ratio, specific loudness calculation, binaural modeling and summation) are discussed later in detail.

2.1 Calibration

Usually recordings are not calibrated. Here it is an aim to model nonlinear effects of human hearing such as shallower masking threshold for higher masker levels, therefor the model has to be calibrated for incoming and outgoing signals to map absolute sound pressure levels to the numerical space of the computer.

2.1.1 Incoming signals

The input calibration is used for analyzing recorded material. A calibration signal $Scal$ has to be recorded. Its RMS¹ level has to be known as $Scal_{dB A, rms, target}$. The calibration is used to map a given input level to the numerical domain of the computer. According to [5] the A-weighted sound pressure level (SPL) of $L_{ListRef}$ is mapped to a numerical level of -18 dB FS. Each reproduction channel is calibrated separately to $L_{ListRef}$, given by equation 1, where n is the number of the reproduction channels.

$$L_{ListRef} = 85 - 10 * \log(n) \text{ dB}(A) \quad (1)$$

The following arithmetics bring the input factor gain that has to be multiplied with the input signal to calibrate.

$$headroom = 18 \text{ dB} \quad (2)$$

$$Scal_{FS, rms, actual} = 20 * \log \sqrt{\frac{1}{N} * \sum_0^N Scal^2} \quad (3)$$

$$Scal_{dB A, rms, actual} = L_{ListRef} + headroom + Scal_{FS, rms, actual} \quad (4)$$

$$gain = 10^{\frac{Scal_{dB A, rms, target} - Scal_{dB A, rms, actual}}{20}} \quad (5)$$

Example:

The headroom is 18 dB, the reference listening level for a mono setup is 85 dB(A). The RMS value of the numerical recorded calibration signal is measured according to equation 3. Its value $Scal_{FS, rms, actual}$ equals -55 dB. This corresponds to a physical level given by equation 4.

¹ root mean square

$$Scal_{dBA,rms,actual} = 85 \text{ dB} + 18 \text{ dB} + (-55 \text{ dB}) = 48 \text{ dB}$$

In the physical domain the level of the calibration signal $Scal_{dBA,rms,target}$ equals 94 dB(A)rms. The input gain for the right input mapping is given by equation 5.

$$gain = 10^{\frac{Scal_{dBA,rms,target} - Scal_{dBA,rms,actual}}{20}} = 10^{\frac{94-48}{20}} = 199.5262$$

$$Sig_{calibrated} = 199.5262 * Sig_{uncalibrated}$$

2.1.2 Outgoing signals

The output calibration is used to reach a desired output level for demonstrating specific sound pressure levels for listening experiments. According to [5] the A-weighted sound pressure level (SPL) at the listening point is given by equation 1. This reference listening level is mapped to -18 dB FS (full scale), which corresponds to a 18 dB headroom.

The following equations bring the gain factor for synchronizing physical and numerical representation.

$$headroom = 18 \text{ dB} \tag{6}$$

$$LFS_{rms,target} = LdBA_{rms,target} - L_{ListRef} - headroom \tag{7}$$

$$LFS_{rms,actual} = \sqrt{\frac{1}{N} * \sum_0^N x^2} \tag{8}$$

$$gain = 10^{\frac{LFS_{rms,target} - LFS_{rms,actual}}{20}} \tag{9}$$

Example:

A signal is created in the numerical space. The desired output level for the signal, named $LdBA_{rms,target}$ equals 47 dB. The listening reference level is 85 dB(A), the headroom is 18 dB. $LFS_{rms,target}$ is derived from equation 7.

$$LFS_{rms,target} = 47 \text{ dB} - 85 \text{ dB} - 18 \text{ dB} = -56 \text{ dB}$$

$LFS_{rms,actual}$ can be measured using equation 8. In MATLAB e.g. 0 dB FS correspond to +/- 1, $LFS_{rms,actual}$ could be measured e.g. as -80 dB, depending of the procedure of which the signal was created. Using equation 9 it is easy to derive the needed gain factor for the correct mapping of the signal.

$$gain = 10^{\frac{-56 \text{ dB} - (-80 \text{ dB})}{20}} = 15.8489$$

$$Sig_{calibrated} = 15.8489 * Sig_{uncalibrated}$$

Calibration of the output amplifier:

In an experimental free field setup, one can create a pink noise signal or a 1 kHz sinusoid signal with e.g. $LdBA_{rms,target} = 94 \text{ dB}^2$, and measure the SPL at the listening point. The output amplifier must than be set for the measuring instrument to display 94 dB(A), rms, slow.

In an experimental phones setup, the output amplifier must be set that the Outer ear- / Phones transfer function (see section 2.2.2) is 0 dB at 1 kHz.

2.2 Outer ear filtering

In general there are some different approaches for modeling the outer ear: An often cited concept is e.g. the MAF/MAP concept. Those frequency curves could be used for modeling the outer ear. MAF stands for "minimum audible field" and corresponds to the absolute hearing threshold in 0° free field representation. The abbreviation MAP stands for "minimum audible pressure" and corresponds to the threshold related to sound pressure at the ear drum. Therefor MAP is likely to be used for headphone modeling.

An extensive but effective way to model the outer ear in an particular experiment situation is to measure its transfer function. The benefit of this method is that one can measure everything effecting the signal between its numerical representation in the computer and the arrival at the ear drum. This takes into account the whole playback setup including the frequency response of the DAC's, the power amplifiers, the speakers, room properties (if wanted) as well as torso reflections, pinna reflections and resonance effects in the ear canal.

2.2.1 Stereo setup

Figure 2 shows the transfer paths from stereo loudspeakers to the listeners' ears. The left speaker's signal L is built by the sum of the left signal part SL and the left masker part ML. The right speaker's signal R is built by the sum of the right signal part SR and the right masker part MR. The propagation paths from the left and the right loudspeaker to the left and right ear are described by the transfer functions $HRTF_{30,L}$, $HRTF_{30,R}$, $HRTF_{330,L}$ and $HRTF_{330,R}$ ³ which are depicted in figure 2. Those function have been measured using a bruel and

² 94 dB SPL at 1 kHz = 1 PaRMS

³ Head Related Transfer Function

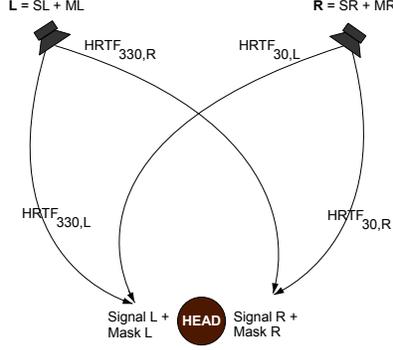


Figure 2: Crosstalk schematic for stereo setup

kjaer dummy head microphone (static head at typical listening position). The used loudspeakers were Genelec 1032 A, with all Dips off (linear response). The input sensitivity 100 dB SPL @ 1 m was adjusted to + 1 dBu⁴.

Figures 3 and 4 show the transfer functions where all room-related effects were excluded to get smoother curves without comb filter notches (free-field conditions, only 64 samples of the impulse response were measured). One can expect symmetrical measurements, but realistic conditions show some variances.

Equations 10, 11, 12 and 13 describe the transfer and crosstalk in a stereo setup, convolving the time signals SL , SR , ML and MR with the impulse responses $HRTF_{330,L}$, $HRTF_{30,L}$, $HRTF_{330,R}$ and $HRTF_{30,R}$.

$$SignalL = HRTF_{330,L} * SL + HRTF_{30,L} * SR \quad (10)$$

$$MaskL = HRTF_{330,L} * ML + HRTF_{30,L} * MR \quad (11)$$

$$SignalR = HRTF_{30,R} * SR + HRTF_{330,R} * SL \quad (12)$$

$$MaskR = HRTF_{30,R} * MR + HRTF_{330,R} * ML \quad (13)$$

2.2.2 Headphones setup

The model structure for headphones setup is simpler than the model structure for stereo setup because there is no noteworthy crosstalk from one ear to the other.

⁴ 0 dBu = 0.775 V

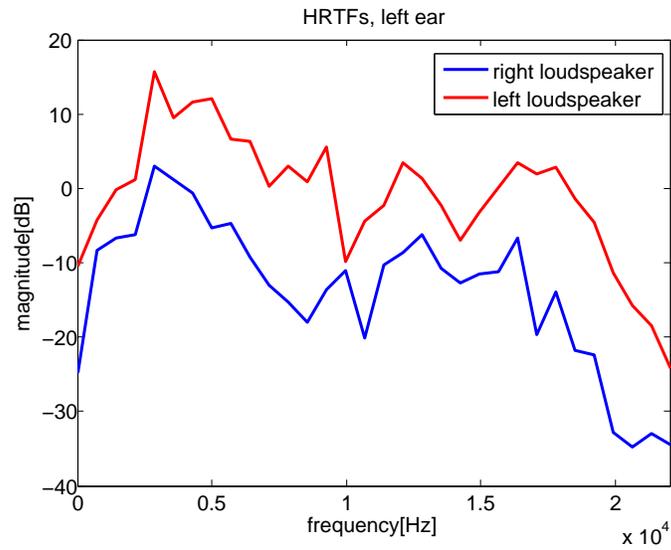


Figure 3: Measured HRTFs, left ear, IR 64 samples

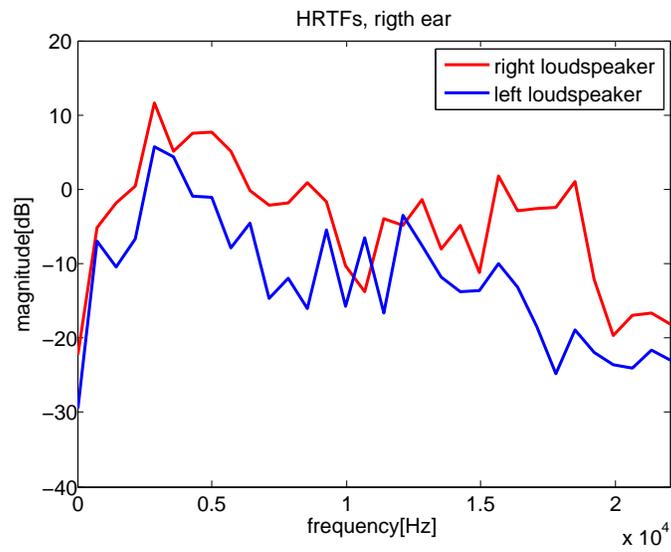


Figure 4: Measured HRTF, right ear, IR 64 samples

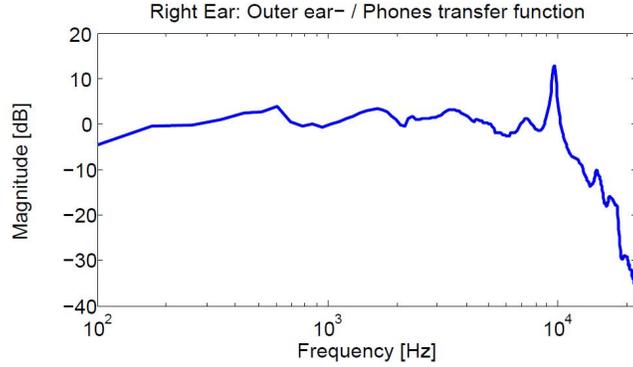


Figure 5: Measured outer ear- / phones transfer function, right ear

$$SignalL = HRTF_L * SL \quad (14)$$

$$MaskL = HRTF_L * ML \quad (15)$$

$$SignalR = HRTF_R * SR \quad (16)$$

$$MaskR = HRTF_R * MR \quad (17)$$

When measuring transfer functions for a headphone setup a stochastic element has been observed: Every time the phones are put off and on again results for higher frequencies differ from previous measures. To compensate those effects five measurements were executed and their absolute transfer functions were averaged. The length of the impulse response was less than 512 samples, so a 512-point DFT⁵ was calculated. No comb filter effects were measured even for extremely high frequency resolution (65536-point DFT).

The measurings have been carried out using a bruel and kjaer dummy head microphone and stereo closed back dynamic headphones, circumaural design: AKG K 271 MK II.

2.3 Middle ear filtering

The transfer function through the middle ear has been taken from [1]. The function describes the transfer from the eardrum through the auditory ossicles

⁵ discrete fourier transformation

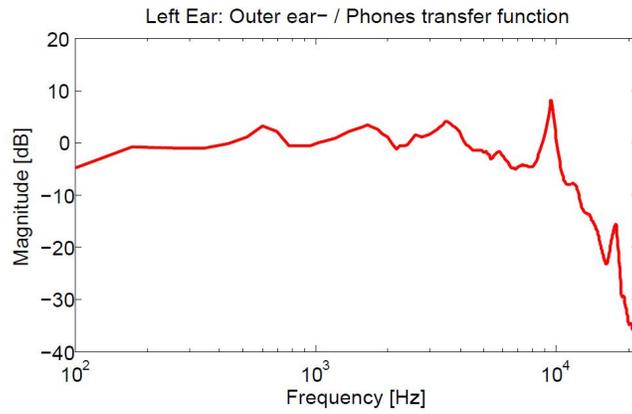


Figure 6: Measured outer ear- / phones transfer function, left ear

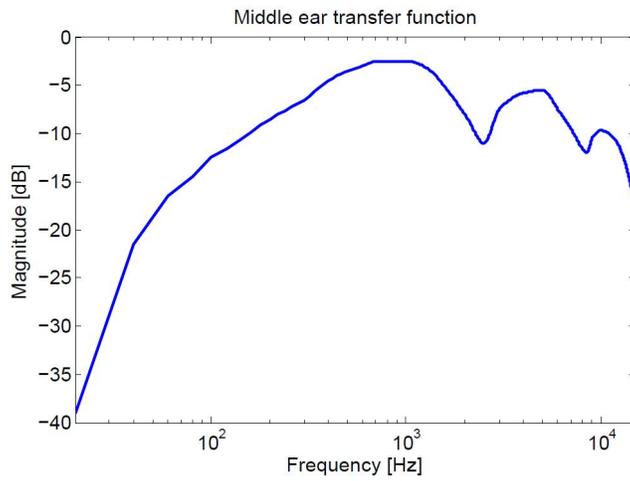


Figure 7: Middle ear transfer function taken from [1]

and the middle ear cavity to the oval window.

2.4 Filter realization

The filters for the phones-setup outer ear and the middle ear can either be realized as linear phase FIR-filters⁶ or as minimum phase IIR-filters⁷. The IIR realization has been chosen because of two reasons: 1. The model is not to be dealing with phase relations of the signals, only magnitude properties will be modeled. 2. An IIR-filter has a smaller order and therefore a smaller delay time (critical in real time applications). The details of the realization are as following: MATLAB's method FIR2 was used to design a 1024th-order FIR-filter. RCEPS calculates the unique minimum-phase sequence that has the same real cepstrum as the sequence of the filter's coefficients. Method PRONY [6] is used to find the coefficients of the IIR-filter having the same transfer function as the FIR-filter.

An IIR-filter with 30 numerator- and 30 denominator- coefficients (i.e. 30th-order) achieves the middle ear's data by less than 1 dB error for frequencies greater than 40 Hz. The two outer ear filters for phones setup have to be 70th-order to achieve the measured data by less than 1.5 dB for frequencies greater than 40 Hz. That is because the measured outer ear data is not as smooth as the middle ear data.

The free-field outer ear filters are simply realized as FIR-filters of 64th order, using the measured impulse response as the filters coefficients.

2.5 Time-frequency analysis

Because the filter bank described in section 2.6 is realized as multiplication in frequency domain, the signal has to be transformed from the time domain to the frequency domain⁸. For audio analyzing there has been observed a problem using the common DFT-algorithm. While frequency resolution and time resolution of DFT-analyzed signals are constant, human hearing works different. According to musical data low frequencies tend to be relative stable in loudness, while the intensity at high frequencies in average changes faster. It is also important that the frequency resolution for low frequencies can somehow image pseudo-discrete spectral bins in a way that those partial tones of the signal are not smeared over the frequency scale too much.

In [2] an analysis algorithm is suggested for the used model. The algorithm is based on 6 parallel dft calculations, each of them using different block lengths and frequency time resolutions. The blocks are hann-windowed each and the 6 independent dfts bring the results for 6 different frequency ranges, achieving the demands described above. Table 1 shows what blocklength was used for what frequencies.

⁶ finite impulse response

⁷ infinite impulse response

⁸ Signal's intensities x^2 are measured to obtain the energy quantity of excitation.

Table 1: Componentual DFT, composition of different block lengths

bin numbers	block length	frequencies
1 - 6	2048 samples	below 80 Hz
7 - 32	1024 samples	80 - 500 Hz
33 - 80	512 samples	500 - 1250 Hz
81 - 162	256 samples	1250 - 2540 Hz
163 - 259	128 samples	2540 - 4050 Hz
260 - 1024	64 samples	4050 - 16000 Hz

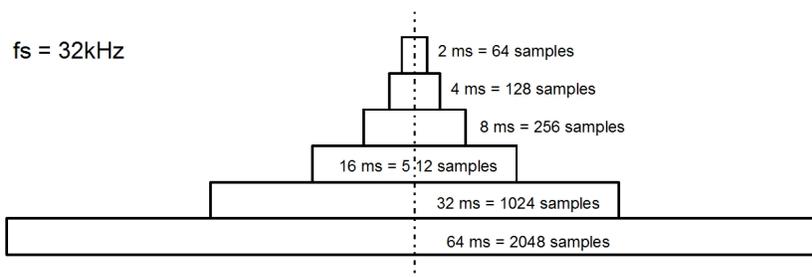


Figure 8: Staged blocklengths for time-domain to frequency-domain transformation

This calculation gives the result spectrum for the time symmetrically centered to the analyze blocks. The blocks are moved forward by 1 ms, so that for every ms a spectrum is calculated.

The following figures show the spectrograms for blocklengths of 2 ms and 64 ms. The last figure shows the componentual spectrogram described above. The analyzed signal was a 1 s piece of a rock song. One can observe that for short blocklengths areas of the same color (i.e. spectral magnitude) tend to spread vertically which testifies for high temporal resolution but poor frequency resolution. For long blocklengths these areas tend to spread horizontally which testifies for poor temporal resolution but high frequency resolution.

Another way to show different frequency resolutions is transform the analyze window to the frequency domain. The figures show the used windows for 64 and for 1024 samples, and their frequency responses.

The relative sidelobe attenuation for a 64 sample window is 42.5 dB and the mainlobe width (-3dB) is 0.039063π [rad/sample] or 625 Hz @ fs = 32 kHz.

The relative sidelobe attenuation for a 1024 sample window is 42.7 dB and

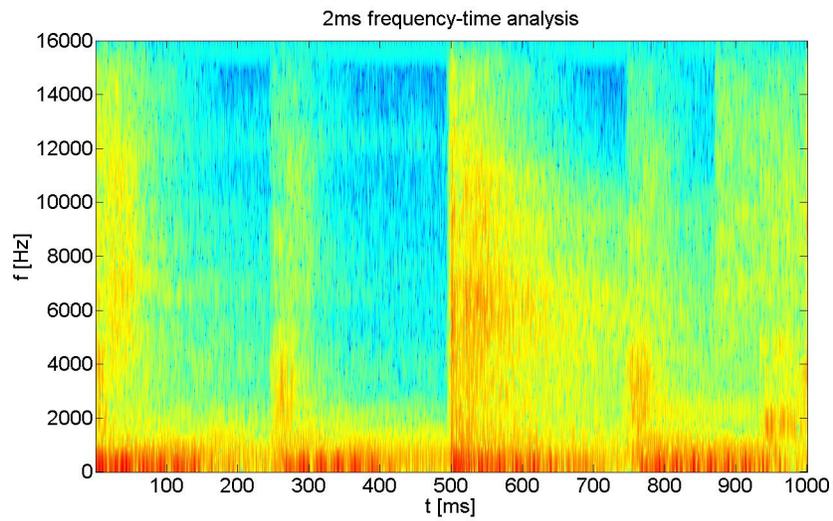


Figure 9: Spectrogram, blocklength = 2 ms

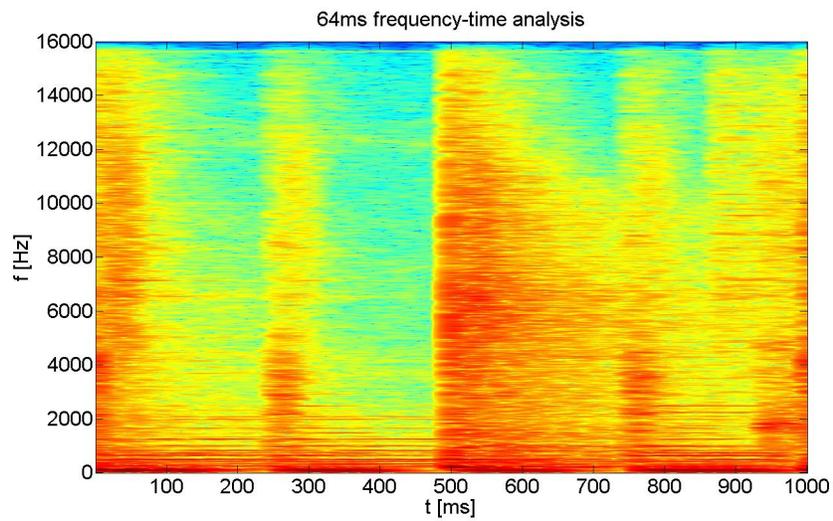


Figure 10: Spectrogram, blocklength = 64 ms

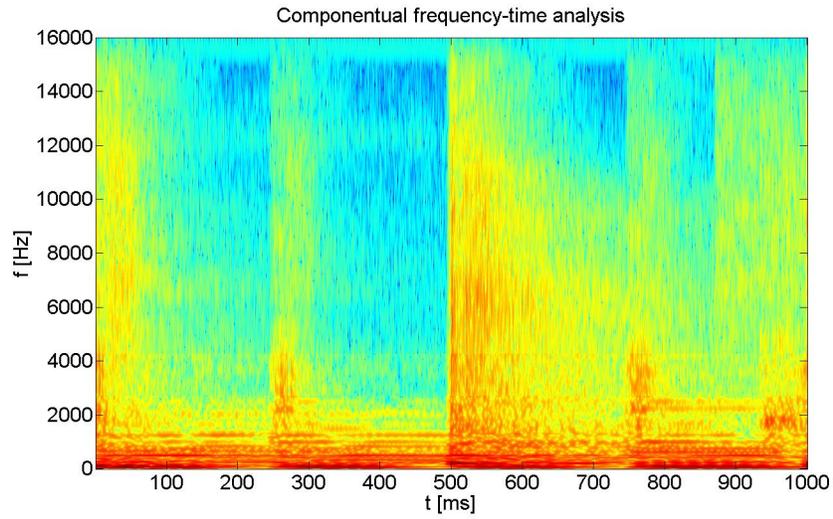


Figure 11: Spectrogram, all used blocklength combined

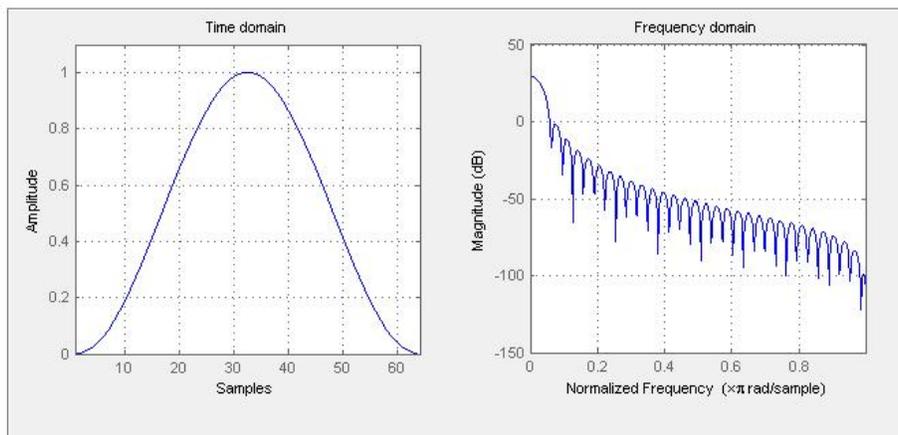


Figure 12: Hann-Window, 64 samples, time-domain versus frequency-domain

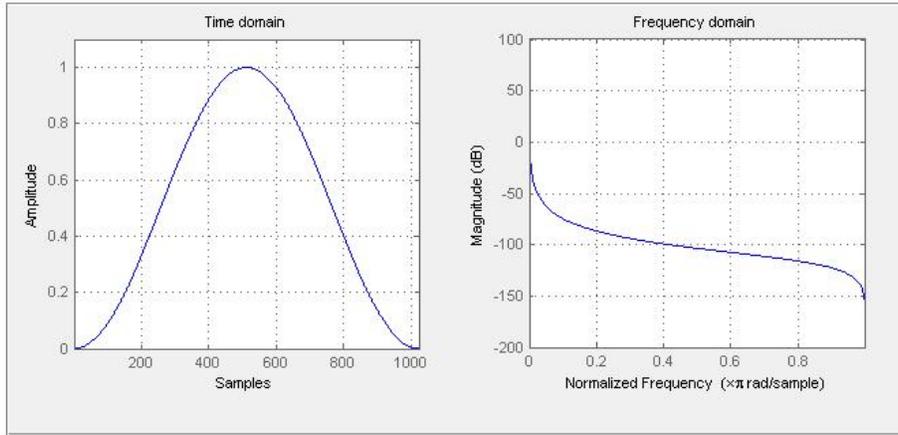


Figure 13: Hann-Window, 1024 samples, time-domain versus frequency-domain

the mainlobe width (-3dB) is 0.0024414π [rad/sample] or 39 Hz @ $f_s = 32$ kHz.

2.6 Cochlear filtering

The prediction if a test signal in a masker environment is audible and how loud the test signal is perceived, needs a calculation of the masking threshold that is derived from the neuronal excitation pattern on the basilar membrane.

Observations from [7] lead to an approach where the basilar membrane is modeled by a filter bank.

1. White noise (constant power over frequency) has a rising masking threshold. The filters of the filter bank therefore must have a growing bandwidth with growing frequency.
2. Tonal maskers have lower masking thresholds than broadband maskers which comes from summation over certain bandwidths (i.e. critical bandwidths). If more than one tonal component is in one critical band, the masking level rises. Therefore it seems a bit raw to simply categorize signals in tonal and non tonal (like done in [8]).
3. If two partials of the same phons-level do lie within a particular frequency range (i.e. a critical band) and therefore excite the same area of the basilar membrane the overall excitation rises by 6 dB, what means the perceived loudness does not double. When two partials of the same phons-level are well apart in frequency loudness does double.

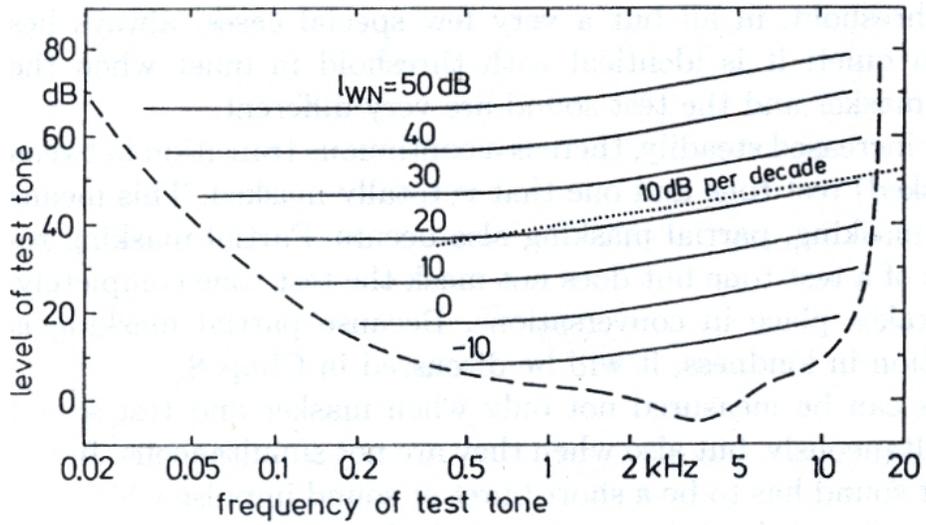


Figure 14: Masking threshold of white noise

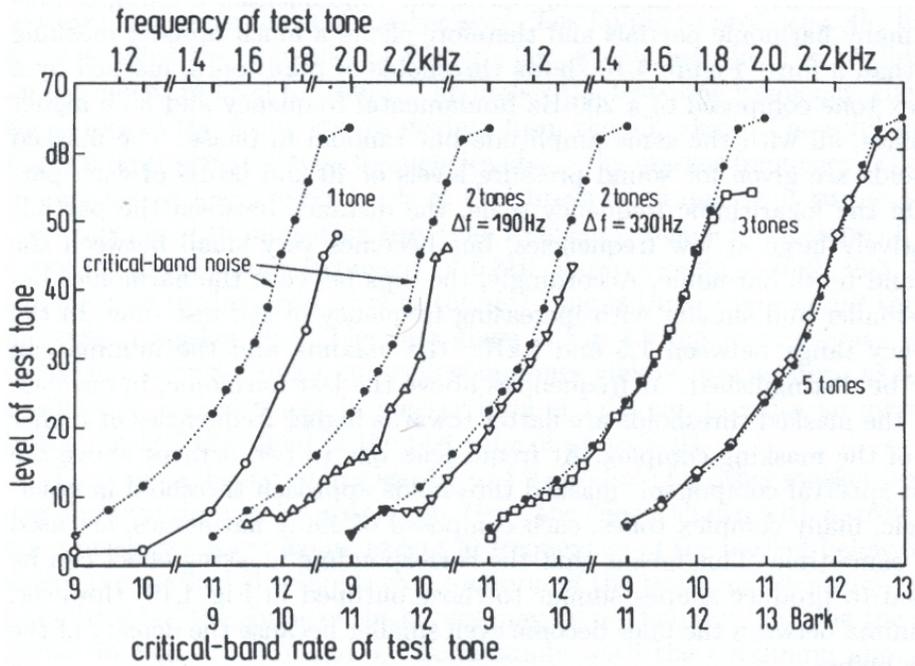


Figure 15: Masking threshold of 1 - 5 sinusoids within one critical band

2.6.1 Details and implementation

Important for the aim of modeling masking thresholds is the spreading function of excitation. As suggested in [9] the frequency behavior of the basilar membrane can be modeled by a filter bank of roex filters, where the output of each filter corresponds to the neuronal excitation at a specific frequency and area of the membrane. The shapes of the filters depend on their center frequencies f_c and the arriving signals' levels $X(f)$. The filters are symmetrical for an input level of 51 dB.

The excitation E is given by:

$$E(fc) = \int_{f=0}^{fs/2} N(f) * W(f, fc) df \quad (18)$$

where $N(f)$ is the spectrum of the masker and $W(f)$ is the filter shape at the excitation's frequency. The integral of this equation models the additive effects of masking (figure 15).

The filter shape $W(g)$ is given by:

$$W(g) = (1 + p * g) * e^{(-p*g)} \quad (19)$$

where g is the normalized deviation from the center of the filter and p is a parameter determining the slope of the filter skirts.

$$g = \frac{|fc - f|}{fc} \quad (20)$$

where fs is the signal's frequency.

The value of p mostly differs for the upper and the lower skirts of the filters giving the parameters p_u and p_l . p_u is assumed to be independent of the input level.

$$p_u = \frac{4 * fc}{ERB} \quad (21)$$

where ERB is the filter's equivalent rectangular bandwidth given by

$$ERB = 24.673 * \left(1 + \frac{4.368 * fc}{1000}\right) \quad (22)$$

where fc is inserted in Hz.

For input levels of 51 dB the filters are symmetrical and therefor

$$p_l(51) = p_u(51) \quad (23)$$

The level dependency of the filters' lower skirts is described by

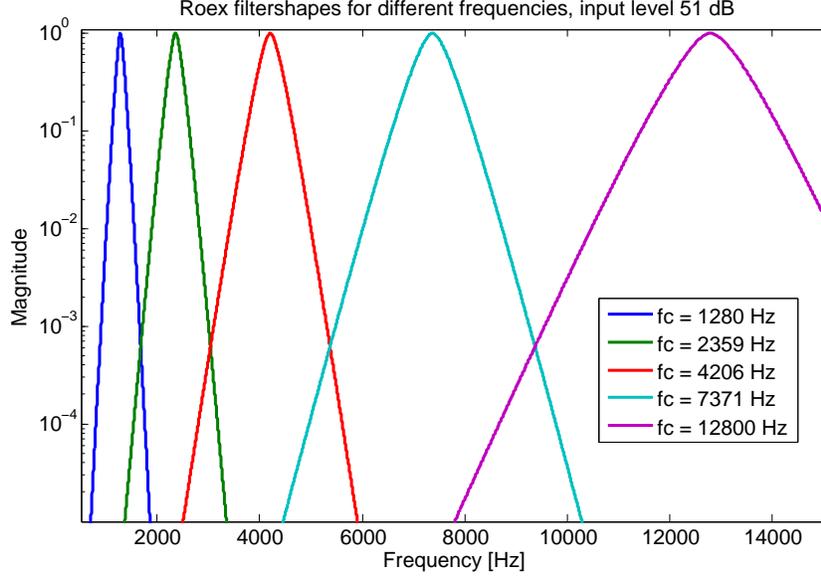


Figure 16: Cochlear modeling: roex filter shapes at 51 dB

$$p_l(X) = p_l(51) - 0.35 * \left(\frac{p_l(51)}{p_l(51, 1k)} \right) * (X - 51) \quad (24)$$

where $p_l(51, 1k)$ is 30.2012922, according to equation 21, 22 and 23. The algorithm was adapted from [10].

Figures 16 and 17 show the roex filter shapes for different frequencies and levels. The filters' lower skirt flattens for higher levels, whereas for lower levels the skirts are sharper. Filters at higher center frequencies are less sharp.

The filterbank has to be implemented in the following way: First the filter-levels have to be determined by filtering the sum of the signal and the masker by symmetrical auditory 51dB-filters. These levels give the values $X(f)$ for equation 24 and give information about the spread of excitation and masking.

Figures 18 and 19 show filter levels for a sinusoid signal and for noise. The filter levels in figure 19 are below 100 dB because 100 dB is the sum level over frequency. The levels are not constant over frequency because of the outer ear filter and the middle ear filter.

Next the signal and the masker run through the now given filterbank separately which brings the excitation patterns E_{sig} and E_{mask} . Let us bring in a perceptual motivated frequency scale "number of ERBs" taken from [1].

$$number\ of\ ERBs = 21.4 * \log(0.00437 * f + 1) \quad (25)$$

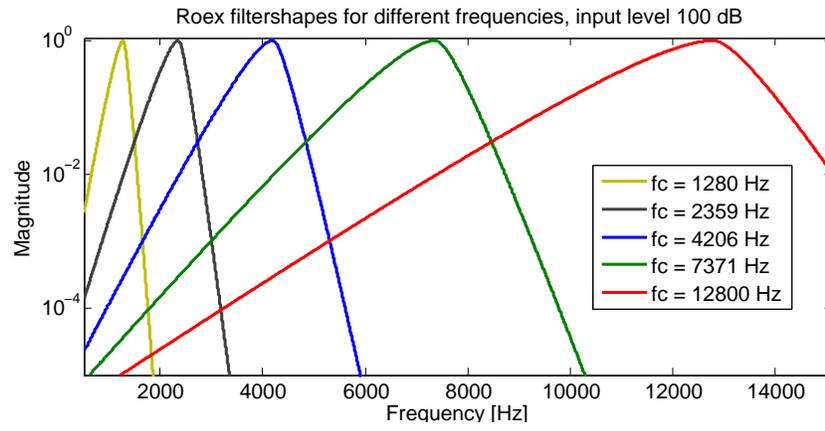


Figure 17: Cochlear modeling: roex filter shapes at 100 dB

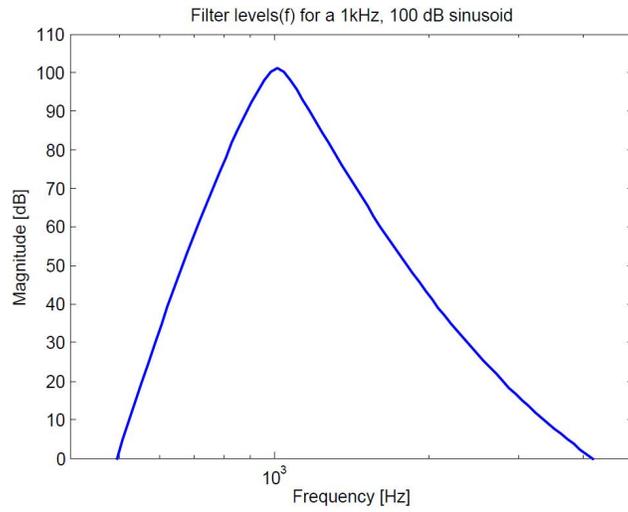


Figure 18: Nonlinear filter bank: filter levels over frequency for a 1kHz, 100 dB sinusoid

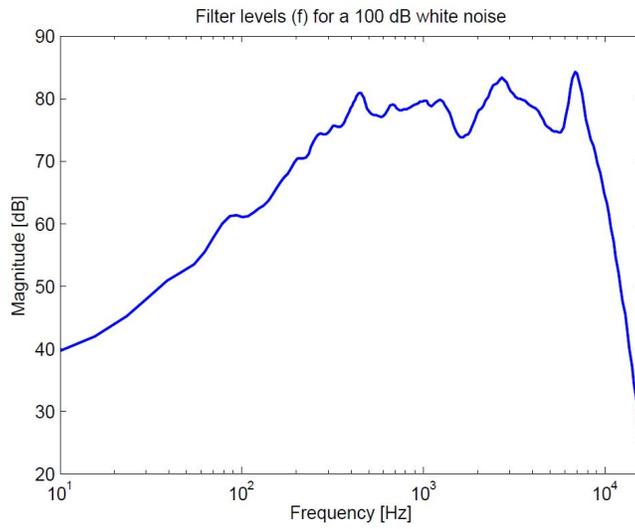


Figure 19: Nonlinear filter bank: filter levels over frequency for a 100 dB white noise

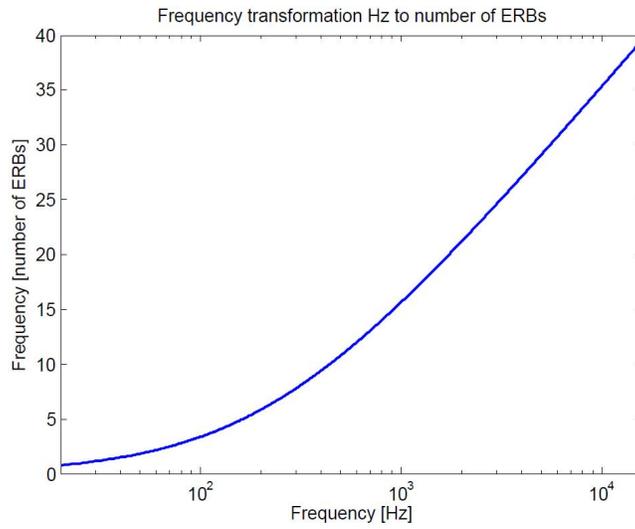


Figure 20: Perceptually motivated frequency scale: "number of ERBs"

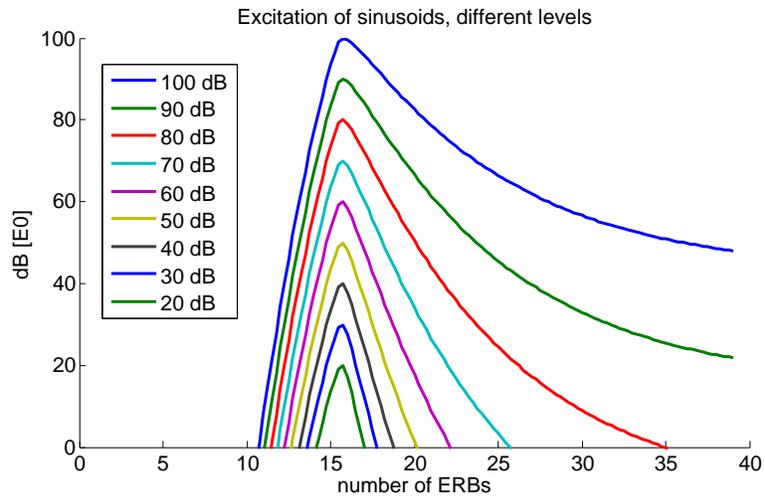


Figure 21: Excitation patterns for 1 kHz sinusoids for the levels 20 to 100 dB in 10 dB steps

Excitation is measured in dB above E0 where E0 is the excitation for a 1 kHz sinusoid at 0 dB SPL.

Figure 22 shows the construction of excitations patterns from auditory filter shapes. The upper diagram shows a 1 kHz sinusoid (dotted line) and the weighting functions of the filters surrounding 1 kHz. The excitation e.g. at 700 Hz is determined by the 700 Hz filter's output (a). Excitation levels at all other frequencies are determined by their corresponding filters each (b - e).

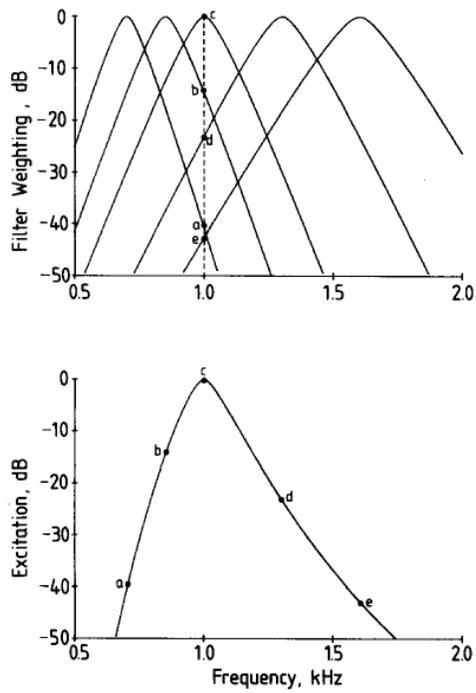


Figure 22: Construction of excitation patterns from auditory filter shapes [20]

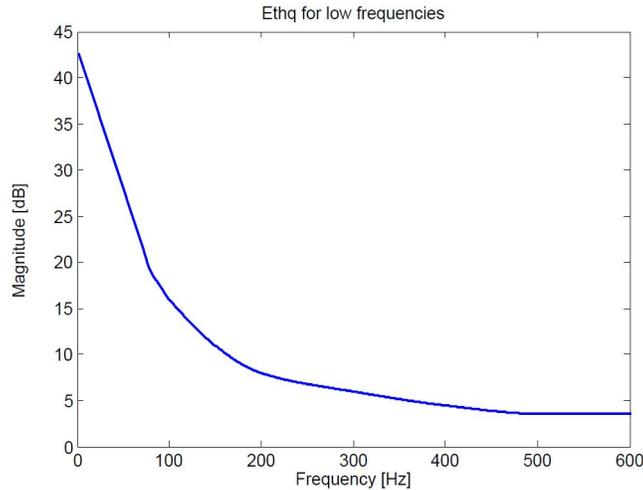


Figure 23: Excitation at threshold in quiet

2.7 Masking threshold

As an extension of equation 18 the excitation of a signal at masking threshold is given by:

$$E_{thr_n} = K * \int_0^{f_s/2} N(f) * W(f) df + E_{thq} \quad (26)$$

where E_{thr_n} stands for excitation at threshold in noise, K (see section 2.7.1) is the excitation-to-mask ratio and E_{thq} is the excitation at threshold in quiet. That means, if a masker becomes by and by quieter and quieter, the masking threshold becomes by and by the absolute hearing threshold.

The excitation at threshold in quiet was also taken from [1], assuming to be constant 3.6 dB for frequencies above 500 Hz and rising for lower frequencies.

Fig. 24 shows analyzing for a flute playing a piano a2. At spectral peaks (i.e. $E \gg E_{thq}$) the red curve is $K(f)$ dBs below the blue curve. Where excitation is lower (i.e. E is slightly greater than E_{thq} or smaller) the threshold is dominated by the threshold in quiet ($f < 8 ERBs$ and $f > 29 ERBs$).

2.7.1 Experiment: adjusting the model

In [1] the K -factor (see section 2.7) is suggested to be constant -3 dB for frequencies above 1 kHz. Below 1 kHz K rises markedly. Its values are shown in figure 25.

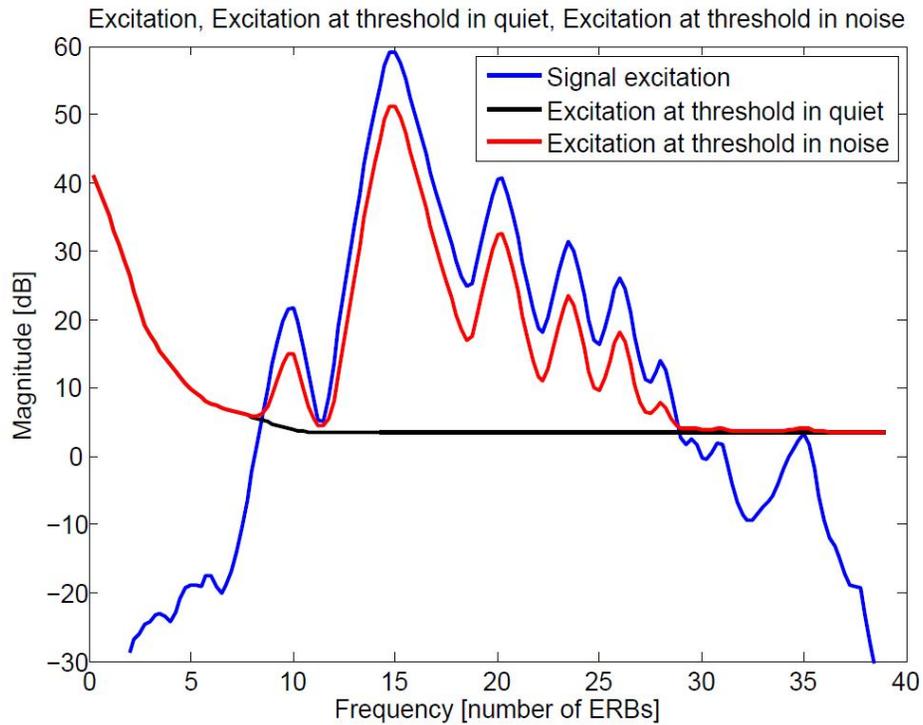


Figure 24: Excitation, Excitation at threshold in quiet, Excitation at threshold in noise

An experiment with 4 ad hoc on hand probands was conducted to adjust the implementation of the masking threshold model. To determine the 79.4 % threshold, according to [11], a transformed up-down method was used (see [12]). The algorithm was adapted from [13]. The masking threshold was measured for a white noise masker of 85 dB SPL, and sinusoid testtones at the frequencies 250, 500, 1000, 2000, 4000 and 8000 Hz. The experiment was separated into 12 runs for each proband (6 frequencies times 2 ears). Before the experiment the probands were asked to adjust the levels of each testtone for each run to be clearly detectable in the masking environment (figure 26). Those levels gave the starting points for the following convergence procedure. A run consists of several trials. In each trial two stimuli in two time intervals were presented to the proband. Both stimuli contained the masker, whereas randomly, only one stimulus contained the testtone. The proband was forced to say if the first or the second interval contains the testtone (figure 27). If the testtone is well below

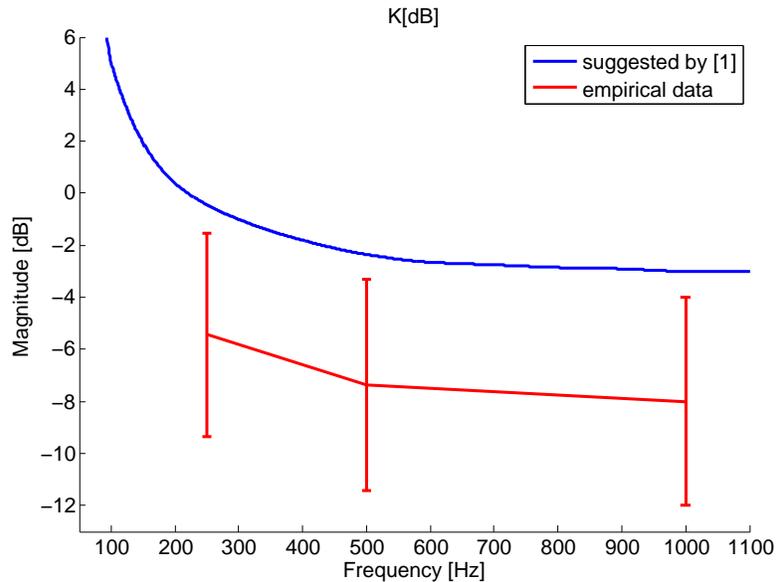


Figure 25: Excitation-to-mask ratio (K-factor), suggested by [1]

masking threshold the probability for a correct answer is 50 % on average. If the testtone is well above masking threshold this probability is nearly 100 %. The convergence procedure works as follows: Everytime the proband gives a wrong guess, the testtone's level is increased by a certain amount. If the proband gives three right answers in a row the testtone's level is decreased by the same amount. This amount is part of tuning the experiment and was chosen to be 6 dB until the first turn point is reached and 3 dB for the rest of the run. Another 4 turnpoints are then awaited. After those first 5 turn points the standard deviation of the testtone's levels is measured before every additional turnpoint. If the standard deviation is below 10 dB or more than 10 turn points were counted for the active run, the run was ended. The measured threshold is given by the mean value of all testtones' levels of one run.

The typical courses of two runs are shown in figure 28. Both runs for the left and the right ear contain 5 turnpoints.

In general it is useful to define any threshold located somewhere around 75 %. For a transformed up-down method it is easy to measure the 79.4 % threshold ($\sqrt[3]{0.5}$).

Figures 29, 30, 31 and 32 show the masking threshold results for each proband separated into left and right ear, the standard deviations of the probands' answers, and the model prediction for the tested scenario built by using the ad-



Figure 26: Adjusting the start levels for measuring the masking threshold. Those start levels have great influence on the following convergence procedure. The proband uses the +, - and play buttons to adjust the level of the testtone to be just audible.

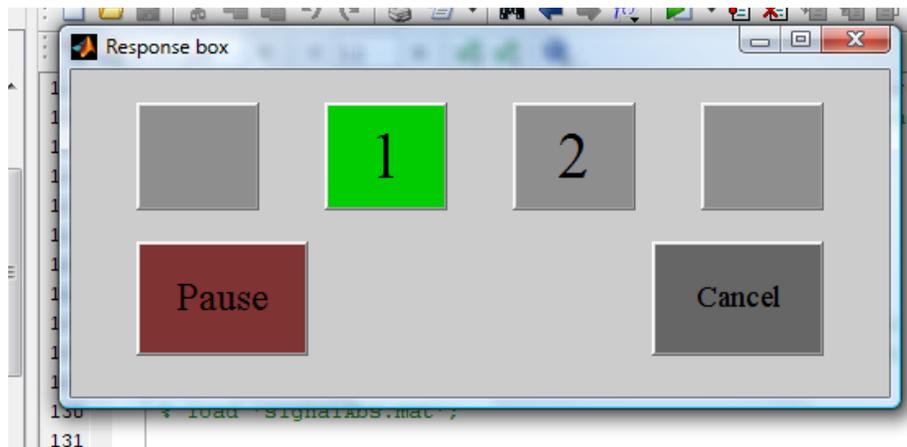


Figure 27: The proband is forced to guess in which time interval a testtone is audible

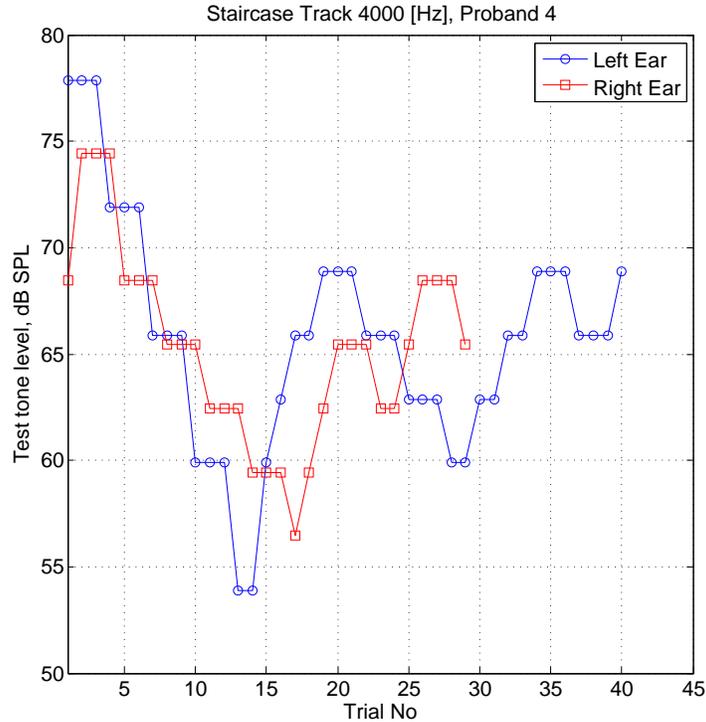


Figure 28: Convergence procedure: typical staircase of presented test tone levels, 1 run consist of several trials. Three correct answers in a row effect a decrease of the test tone level, one false answer effects an increase of the test tone level.

Table 2: Statistics of the measurement: mean, median and standard deviation

f	250 Hz	500 Hz	1 kHz	2 kHz	4 kHz	8 kHz
mean	-33.5 dB	-34.0 dB	-33.4 dB	-33.5 dB	-33.4 dB	-32.7 dB
median	-33 dB					
std	7.9 dB	8.1 dB	8.0 dB	8.2 dB	7.6 dB	7.6 dB

justing factor k proposed by [1].

Table 2 shows the statistics for the results of the convergence procedures. The mean and the median of each frequency lie close together, therefore the quality of the data is good.

Figure 33 shows the averaged empirical data for all probands and all ears, as well as the model prediction derived from Moore's k -factor. It was found, that measured thresholds lie markedly below the model prediction. The greatest deviation is about 7 dB, and measurement data is below predictions for all frequencies. If the k -factor was reduced by 5 dB (black line) the compliance of the prediction and the measurement is much better. The greatest deviation is

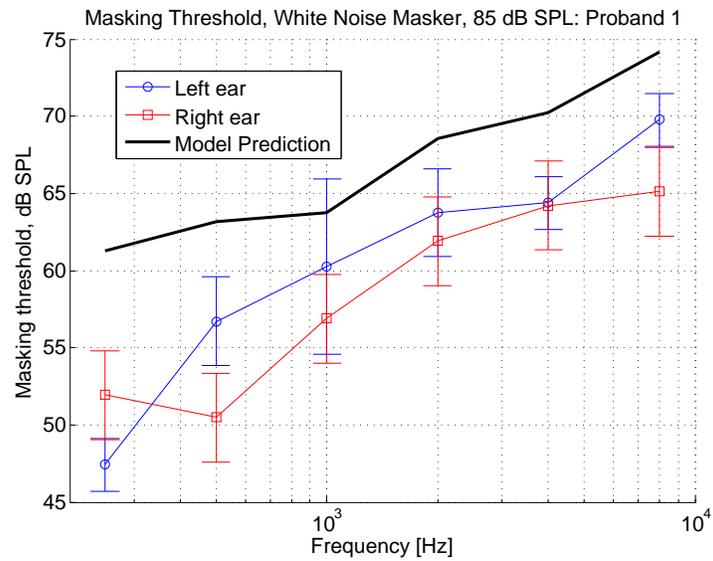


Figure 29: Measured masking threshold for a white noise masker, 85 dB(SPL), Proband 1

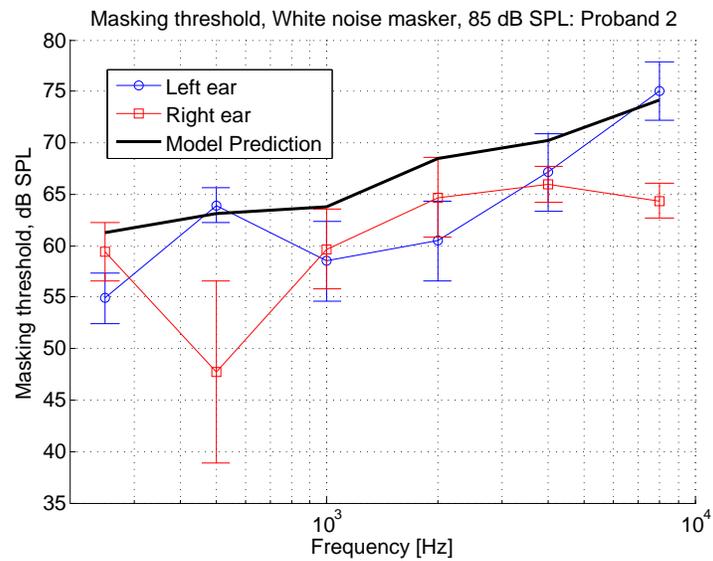


Figure 30: Measured masking threshold for a white noise masker, 85 dB(SPL), Proband 2

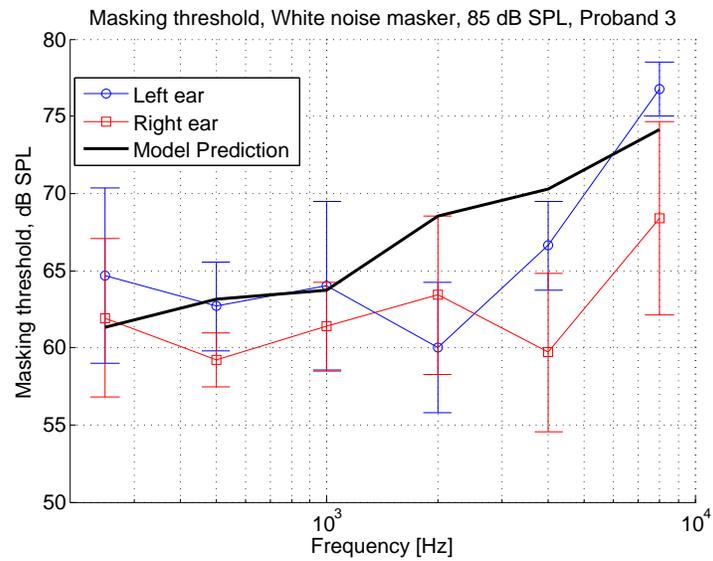


Figure 31: Measured masking threshold for a white noise masker, 85 dB(SPL), Proband 3

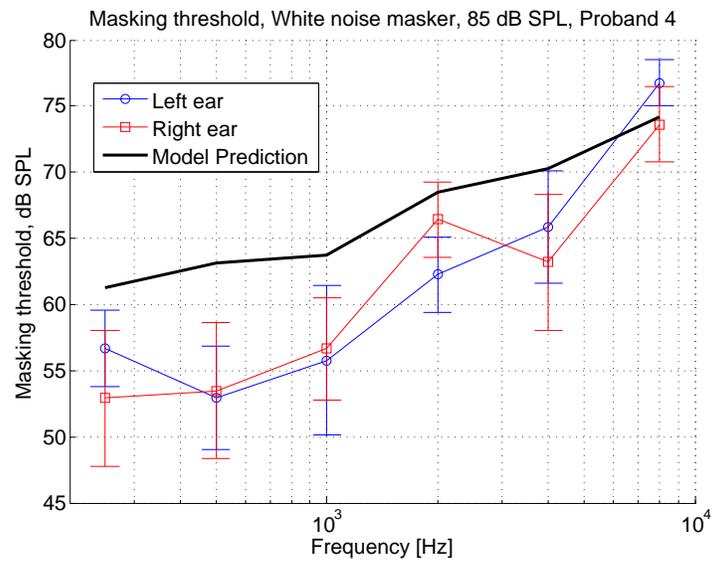


Figure 32: Measured masking threshold for a white noise masker, 85 dB(SPL), Proband 4

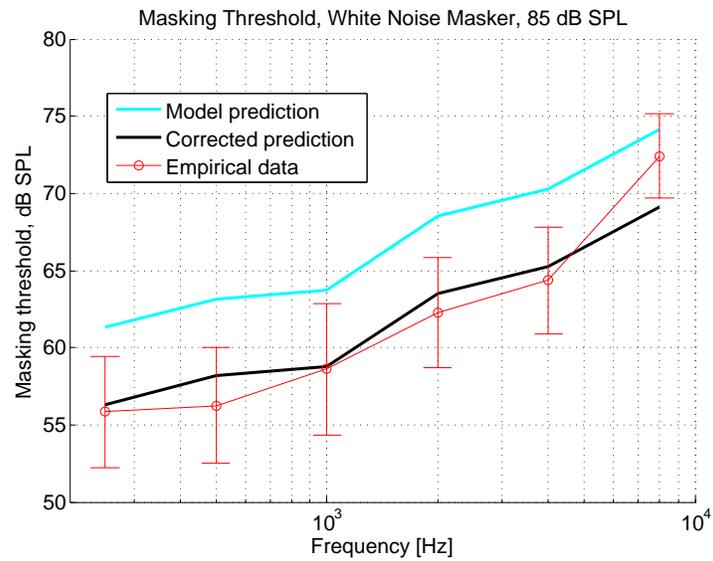


Figure 33: Averaged result versus model prediction

than about 3.4 dB at 8 kHz.

Figure 34 shows the same experiment for just one proband and a narrow band noise masker centered at 410 Hz. Measured frequencies were 150, 250, 410, 650 and 1000 Hz. Again, all predictions are too high, except for 150 Hz. This could be a result of imperfect experiment conditions (i.e. low frequency ambient noise). All other deviations are assumed to be proband's variance.

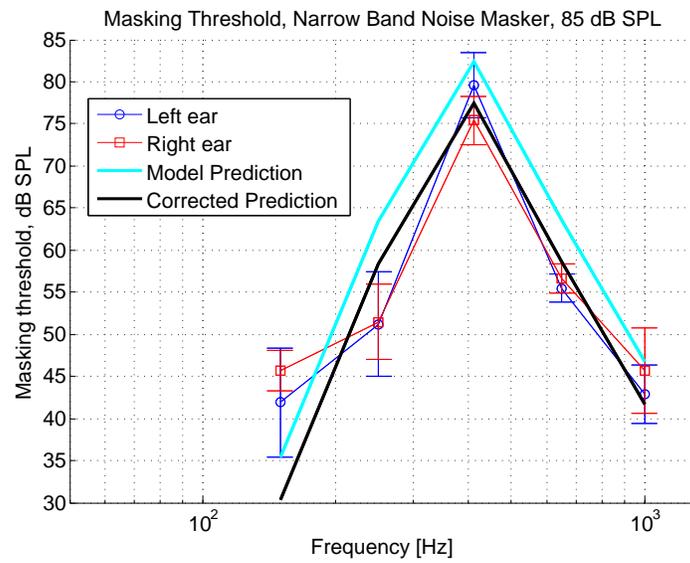


Figure 34: Measured masking threshold for a narrow band noise masker, 85 dB(SPL)

2.8 Loudness of steady sounds in noise

This section describes how perceived loudness is derived from excitation patterns. At first perception of steady sounds is described. As further development time varying sounds in time varying noise will be covered later on in section 2.9.

2.8.1 Partial or specific loudness

Predicting perceived loudness is not a linear procedure. It is effected by the spectral shape and the absolute level of a signal. Therefor calculations have to be done for a number of frequency bands (see section 2.6) in a calibrated environment (see section 2.1).

”The physical characteristics of sound that mediate the perception of timbre include spectrum and envelope. Timbre is also known in psychoacoustics as tone quality or tone color. For example, timbre is what, with a little practice, people use to distinguish the saxophone from the trumpet in a jazz group, even if both instruments are playing notes at the same pitch and loudness.” [21]

To objectively judge about partly masked (i.e. mixed) instruments it is nearby to use the specific loudness (i.e. loudness versus frequency) derived from the described model, because it can predict what spectral parts of the instrument are audible at what time (corresponds to spectrum and envelope).

Therefor an important indicator for timbre perception is partial loudness. Those loudness patterns are derived from excitation patterns and show which parts of a signal’s spectrum have what share of the signal’s overall loudness. Its unity is [sone/bark] and its derivation is shown in [1].

We have to distinguish between 4 cases:

1. $E_{sig} \geq E_{thrn} \wedge E_{sig} + E_{mask} \leq 10^{10}$:
(Signal is above masking threshold, masker and signal do not exceed 100 dB)
2. $E_{sig} < E_{thrn} \wedge E_{sig} + E_{mask} \leq 10^{10}$:
(Signal is below masking threshold, masker and signal do not exceed 100 dB)
3. $E_{sig} \geq E_{thrn} \wedge E_{sig} + E_{mask} > 10^{10}$:
(Signal is above masking threshold, masker and signal do exceed 100 dB)
4. $E_{sig} < E_{thrn} \wedge E_{sig} + E_{mask} > 10^{10}$:
(Signal is below masking threshold, masker and signal do exceed 100 dB)

The inequalities of E_{sig} and E_{thrn} describe if a signal in noisy environment is above or below masking threshold. The inequalities of $E_{sig} + E_{mask}$ and 10^{10} describe if the sum of masker and signal is above or below 100 dB⁹. For those

⁹ Excitation is an energy quantity, therefore $10^{10} = 100 \text{ dB}$

four cases four different perceptual behaviors have been investigated and therefor four different equations for partial loudness have been formed.

Case 1

$$\begin{aligned}
N'_{sig} = & C * \{[(E_{sig} + E_{mask}) * G + A]^\alpha - A^\alpha\} \\
& - C * \{[(E_{mask} * (1 + K) + E_{thq}) * G + A]^\alpha \\
& - (E_{thq} * G + A)^\alpha\} * \left(\frac{E_{thrn}}{E_{sig}}\right)^{0.3}
\end{aligned} \tag{27}$$

Case 2

$$\begin{aligned}
N'_{sig} = & C * \left(\frac{2 * E_{sig}}{E_{sig} + E_{thrn}}\right)^{1.5} \\
& * \left\{ \frac{(E_{thq} * G + A)^\alpha - A^\alpha}{[(E_{mask} * (1 + K) + E_{thq}) * G + A]^\alpha - (E_{mask} * G + A)^\alpha} \right\} \\
& * [(E_{sig} + E_{mask}) * G + A]^\alpha - (E_{mask} * G + A)^\alpha
\end{aligned} \tag{28}$$

Case 3

$$\begin{aligned}
N'_{sig} = & C_2 * (E_{sig} + E_{mask})^{0.5} \\
& - C_2 * \{[(1 + K) * E_{mask} + E_{thq}]^{0.5} \\
& - (E_{thq} * G + A)^\alpha + A^\alpha\} \left(\frac{E_{thrn}}{E_{sig}}\right)^{0.3}
\end{aligned} \tag{29}$$

Case 4

$$\begin{aligned}
N'_{sig} = & C * \left(\frac{2 * E_{sig}}{E_{sig} + E_{thrn}}\right)^{1.5} \\
& * \left\{ \frac{(E_{thq} * G + A)^\alpha - A^\alpha}{[E_{mask} * (1 + K) + E_{thq}]^{0.5} - (E_{mask})^{0.5}} \right\} \\
& * [(E_{sig} + E_{mask})^{0.5} - (E_{mask})^{0.5}]
\end{aligned} \tag{30}$$

N'_{sig} is the partial loudness of a signal, the low-level gain of the cochlear amplifier G is given by figure 35, the slope design factor A is given by figure 36 and compression exponent α is given by figure 37.

The constant C is chosen so that the peak specific loudness of a 1-kHz sinusoid of 40 dB SPL is predicted as 0.2313 sone/bark for a binaural model ($C = 0.047$)

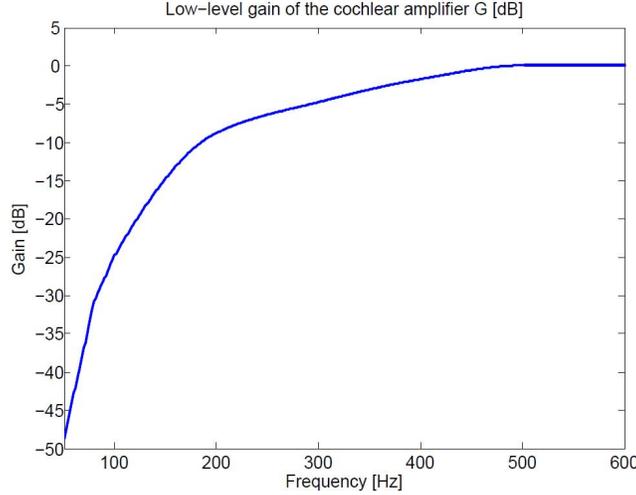


Figure 35: low-level gain of the cochlear amplifier G [dB]

or 0.1732 sone/bark for a monaural model ($C = 0.0352$). The constant C_2 is chosen so that eqs. 27 and 29 are equal at $E_{sig} = 10^{10}$ ($C_2 = C/1040000$). These target values are read out from the model's implementation.

Figure 2.8.1 shows the relation between a signal's excitation level and its partial loudness without a masker. One can see the frequency dependence in different colored curves. Every excitation at threshold in quiet corresponds to a frequency (see section 23). Here the pairs are: 3.6 dB and 500 Hz and above, 6.3 dB and 253 Hz, 14.5 dB and 108 Hz, 20.2 dB and 74 Hz, 26.2 dB and 52 Hz. The black horizontal line shows specific loudness at hearing threshold at 0.00537 [sone/bark] that is subtended by the curves at their threshold excitation levels. The curves are divided into three sections: One section for excitation levels above 100 dB or specific loudness above 4.62 sone/bark, one section for signals below threshold (i.e. $N' \leq 0.00537$ [sone/bark]), and one section for moderate levels. The loudness at threshold of hearing is not defined as 0 sone because of common threshold definition. In a 2AFC-Experiment the hearing threshold is defined as the level for which 75 percent of the probands' answers are correct. Therefore can a signal's loudness at hearing threshold not be 0 but simply "hard to recognize".

When a masker is turned on the signal's loudness is damped. Figure 2.8.1 shows signal excitation level versus signal specific loudness for 1 kHz for different masker excitation levels. The curve $E_{mask} = -\infty$ equals the leftmost curve in the figure above. For other masker excitation levels the signal excitation has to reach a certain level to be audible i.e. the intersections of the curves with

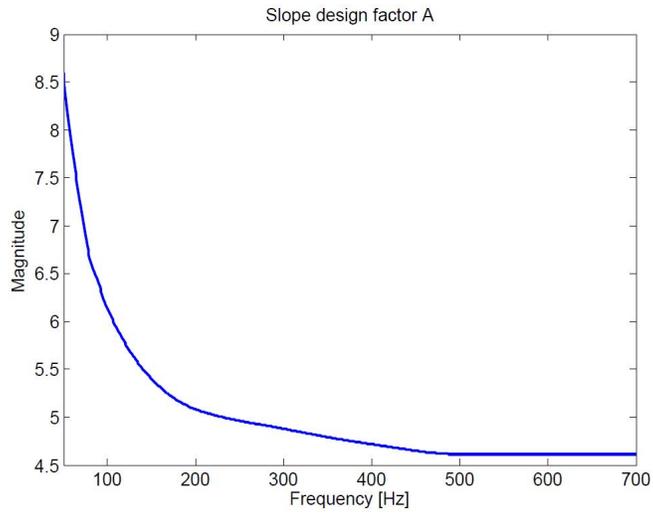


Figure 36: Slope design factor A

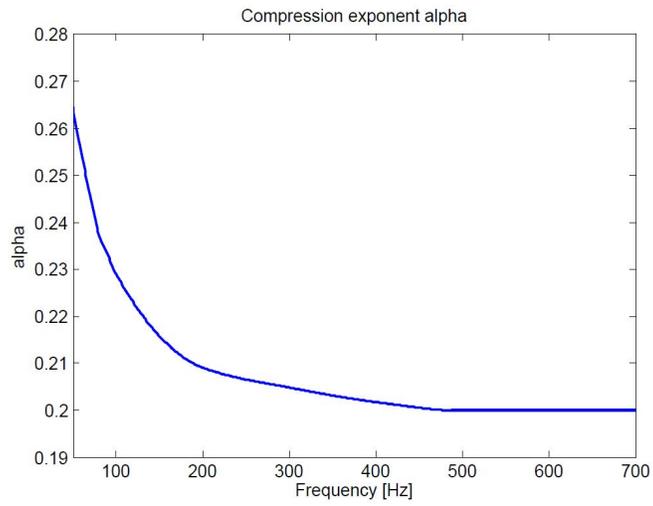


Figure 37: Compression exponent alpha

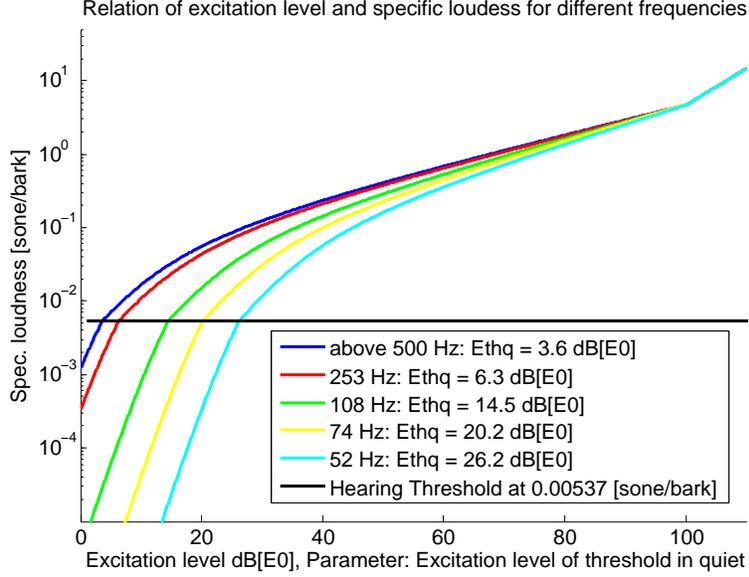


Figure 38: Relation of excitation level and specific loudness for different frequencies

the horizontal line of hearing threshold, that are $K(f = 1000Hz)$ dB below E_{mask} . As E_{sig} rises higher above threshold the signal's loudness convergates to its undamped value.

2.8.2 Overall or total loudness

The overall or total loudness in sone is calculated by summing the specific loudness for all frequencies (eq. 31). Specific loudness can be calculated in 0.25-ERB, 0.1-ERB or other intervals to find a trade off between accuracy and computing time (computational complexity grows with growing number of filters). This sum has to be divided by a factor so that the overall loudness of a 1-kHz sinusoid at 40 dB monaural is 0.75 sone for a monaural model or 1 sone for a binaural model [14].

$$N = \int_0^{f_s/2} N' / C_3 \quad (31)$$

where C_3 was 2.373.

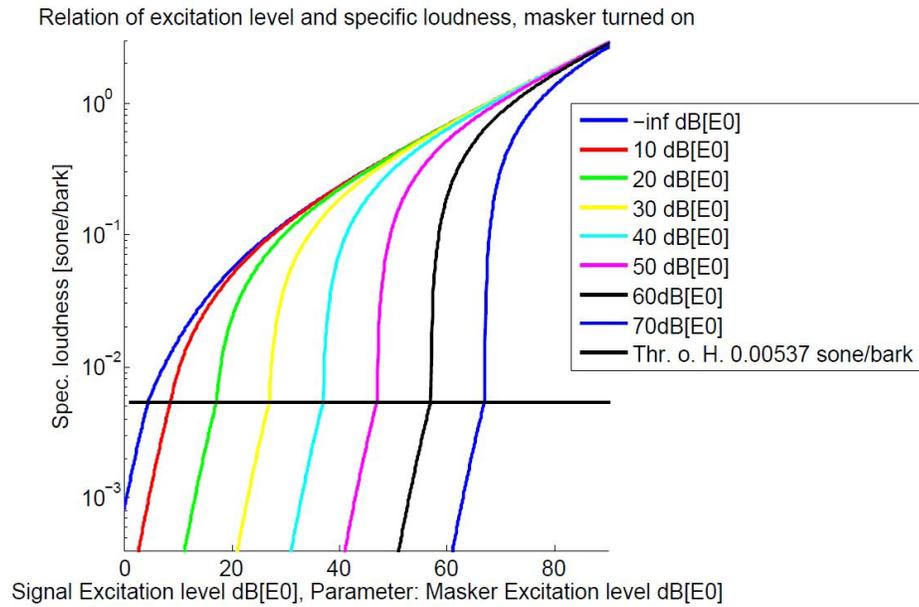


Figure 39: Relation of excitation level and specific loudness, masker turned on

Figure 2.8.2 shows the predicted overall loudness of a 1kHz sinusoid at 1 kHz 40 dB SPL for a binaural model (= 40 phons). The loudness of 40 phons is defined to be 1 sone. A rise by 10 phons doubles the perceived loudness, a fall by 10 phons halves the perceived loudness. At low levels loudness falls quicker to reach 0.003 sone at 2 phons (hearing threshold). At high levels loudness rises faster to fit empirical data [1].

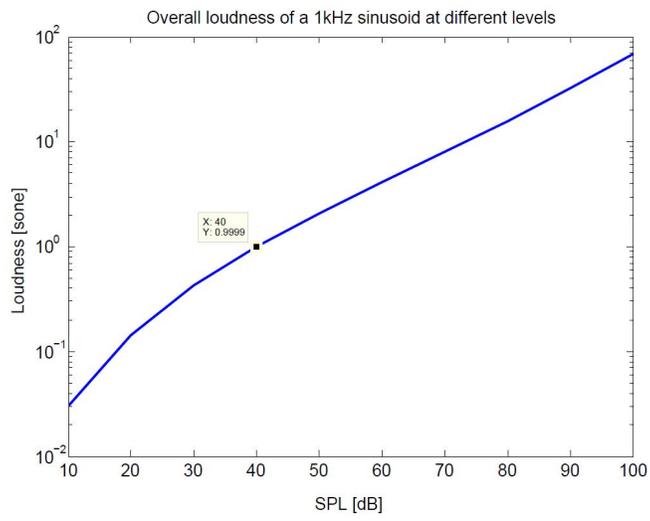


Figure 40: Overall loudness of a 1kHz sinusoid at different levels

2.9 Loudness of timevarying sounds in timevarying noise [2]

If a real musical signal or any other time-varying signal is analyzed, loudness is calculated for every ms (see section 2.5). The so calculated loudness is called instantaneous loudness. However human hearing is based on perceptual time slices that are longer than 1 ms. Therefore two new variables are defined: short-term loudness and long-term loudness. Those can be calculated by integrating instantaneous loudness with different time constants. In speech recognition for example short-term loudness can measure the loudness of an isolated syllable or a single word and long-term loudness can measure the loudness of whole sentences or text passages. In musical perception short-term loudness can be used for short musical notes and long-term loudness can be used for measuring the loudness of a musical phrases or whole pieces.

2.9.1 Calculation of short-term loudness

When $N_{st}[n] > N_{st}[n-1]$, corresponding to an attack of loudness, the difference equation of Short-term loudness $N_{st}[n]$ is

$$N_{st}[n] = \tau_{as} * N[n] + (1 - \tau_{as}) * N_{st}[n-1] \quad (32)$$

where $N[n]$ is the instantaneous loudness in time slice n and τ_{as} (time constant of attack) is given by

$$\tau_{as} = 1 - e^{-T_i/T_{as}} \quad (33)$$

where T_i is the time interval between two successive values of the instantaneous loudness N (1 ms in this case) and T_{as} is a time constant chosen to be 0.0217 s.

When $N_{st}[n] \leq N_{st}[n-1]$, corresponding to an release of loudness, a similar difference equation is used, with a different time constant for release: $T_{rs} = 0.0495$ s.

$$N_{st}[n] = \tau_{rs} * N[n] + (1 - \tau_{rs}) * N_{st}[n-1] \quad (34)$$

$$\tau_{rs} = 1 - e^{-T_i/T_{rs}} \quad (35)$$

T_{rs} being greater than T_{as} means that perceived short-term loudness can faster be built up than forgotten. In figure 41 one can see the analysis of a white noise burst, switched on and off immediately.

2.9.2 Calculation of long-term loudness

The equations for long-term loudness $N_{lt}[n]$ are

$$N_{lt}[n] = \tau_{al} * N[n] + (1 - \tau_{al}) * N_{lt}[n - 1] \quad (36)$$

for times of attack ($N_{lt}[n] > N_{lt}[n - 1]$) and

$$N_{lt}[n] = \tau_{rl} * N[n] + (1 - \tau_{rl}) * N_{lt}[n - 1] \quad (37)$$

for times of release ($N_{lt}[n] \leq N_{lt}[n - 1]$)

with

$$\tau_{al} = 1 - e^{-T_i/T_{al}} \quad (38)$$

$$\tau_{rl} = 1 - e^{-T_i/T_{rl}} \quad (39)$$

and

$$T_{al} = 0.0995s \quad (40)$$

$$T_{rl} = 1.9995s \quad (41)$$

Figure 41 shows the analysis of a white noise burst, switched on and off immediately.

2.9.3 Instantaneous partial loudness IPL and short-term partial loudness STPL

As suggested in [11] one can also apply temporal smearing to partial loudness patterns to achieve a greater benefit for temporal timbre analysis. The original partial loudness patterns, updated every ms, are then called IPL (instantaneous partial loudness) and the smeared patterns using short-term time constants are called STPL (short-term partial loudness). One can also define LTPL (long-term partial loudness), that is not defined in [11], using the corresponding long-term time constants.

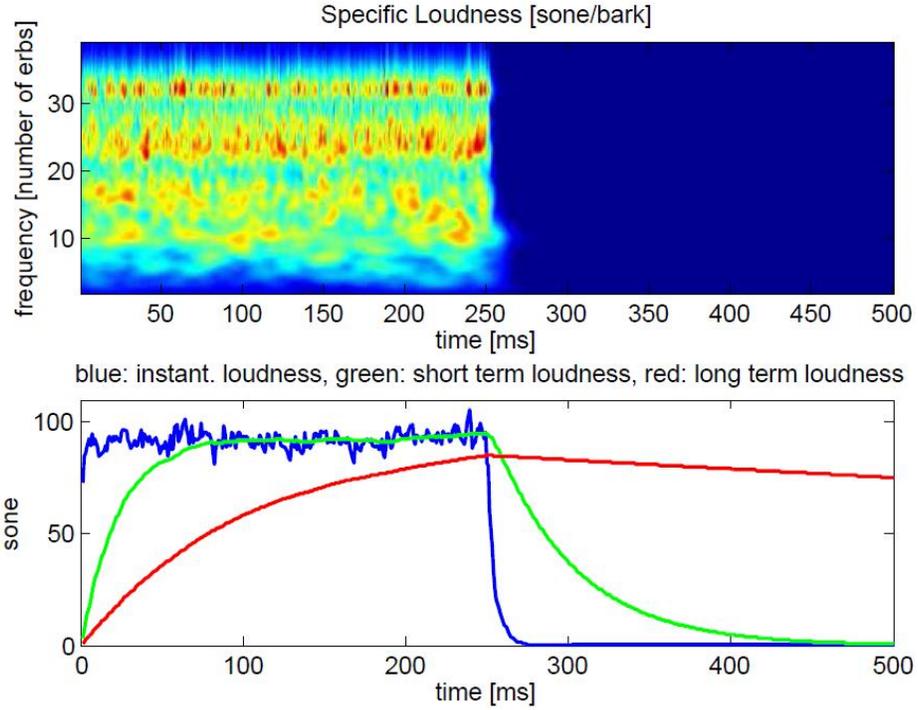


Figure 41: Specific loudness, instant. loudness, short-term loudness, long-term loudness

2.10 Binaural summation of loudness

In order to predict masking relations of panned maskers and signals it is necessary to use a model of binaural loudness. According to recent empirical data [14], summation for loudness across ears was found to be less-than-perfect. That means if two identical sounds are presented first monaurally and then binaurally, perceived binaural loudness is less than twice the monaural loudness. If the left and the right signal are different in spectral appearance, the inhibition of loudness summation is less, depending on specific loudness inhibition tuning functions. This tuning of inhibition is quite broadband.

The tuning function W is given by

$$W(g) = e^{-B \cdot g^2} \quad (42)$$

where B is a constant set to 0.08, g is a normalized frequency variable given by

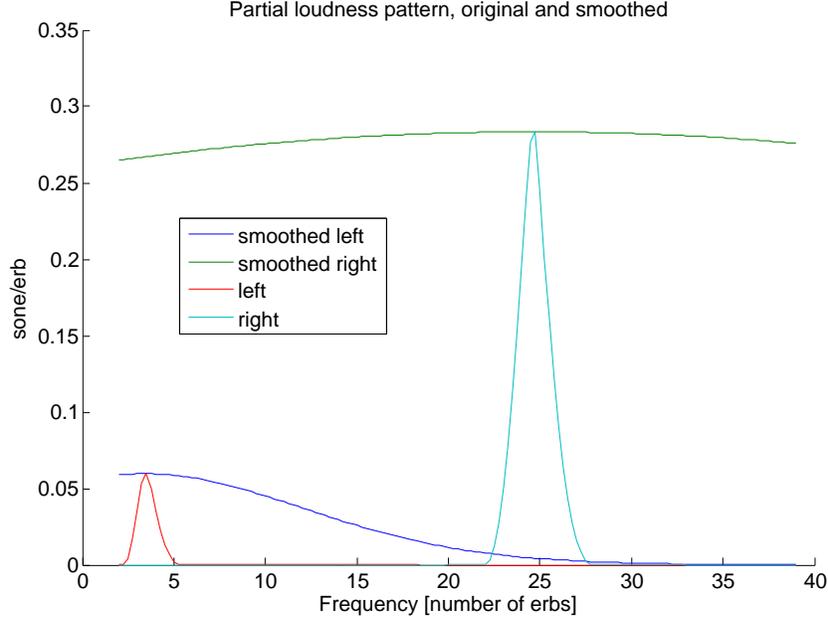


Figure 42: Partial loudness patterns, original and smoothed, using a gaussian spreading function

$$g = \frac{\|E - \Delta E\|}{E} \quad (43)$$

and E equals the frequency unity *number of ERBs* given by equation 25.

The specific loudness patterns of both ears are convolved with the weighting function W . Figure 42 shows the patterns for two sinusoids of different frequencies. One can see that at lower frequency the tuning function is not as broad as at high frequencies.

$$S_L(g)_{smoothed} = N'_L(g) * W(g) \quad (44)$$

$$S_R(g)_{smoothed} = N'_R(g) * W(g) \quad (45)$$

with specific loudness patterns of left and right ear $N'_L(g)$ and $N'_R(g)$.

The left ear signal causes an inhibition at the right ear and vice versa. Therefore the factors $INHIB_L(E)$ and $INHIB_R(E)$ are brought in and depend on the quotient of the smoothed loudness patterns.

$$INHIB_L(g) = 2/[1 + sech(S_R(g)_{smoothed}/S_L(g)_{smoothed})^p] \quad (46)$$

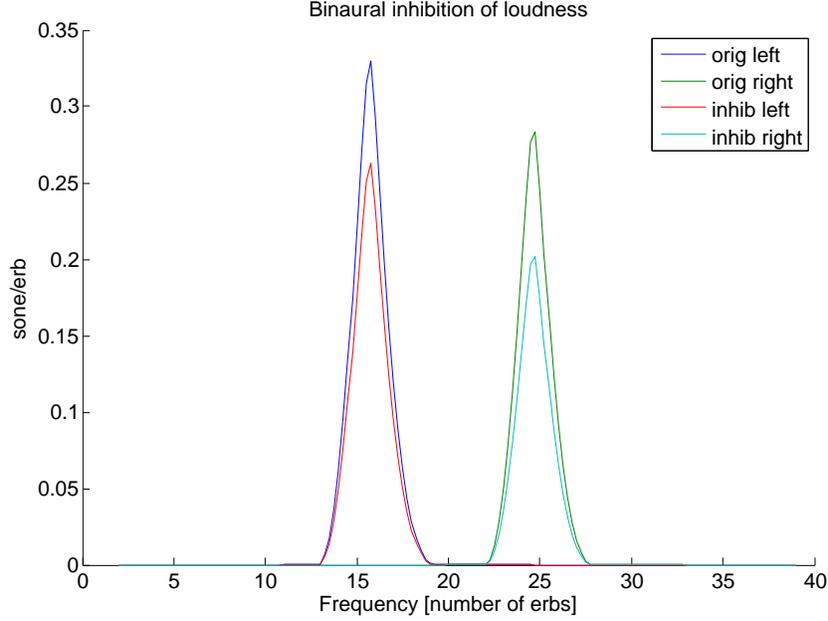


Figure 43: Binaural loudness: original specific loudness versus inhibited specific loudness

$$INHIB_g(E) = 2/[1 + \text{sech}(S_L(g)_{smoothed}/S_R(g)_{smoothed})^P] \quad (47)$$

where *sech* is the hyperbolic secant.

$$\text{sech}(x) = \frac{2}{e^x + e^{-x}} \quad (48)$$

The original loudness patterns are then multiplied with the inhibition factors.

$$N'_L \text{binaural}(g) = N'_L(g) * INHIB_L(g) \quad (49)$$

$$N'_R \text{binaural}(g) = N'_R(g) * INHIB_R(g) \quad (50)$$

Fig. 43 shows original and damped loudness patterns.

The total loudness is than

$$N_{\text{binaural}} = \left(\int_0^{fs/2} N'_R \text{binaural}(g) + \int_0^{fs/2} N'_L \text{binaural}(g) \right) / C_4 \quad (51)$$

where C_4 is chosen so the overall loudness of two 1-kHz sinusoids at diotic representation is 1 sone (2.2885, read out from implementation of the model).

2.10.1 Binaural unmasking for two sound sources

To predict binaural unmasking the models of papers [1] and [14] will be merged.

2.10.1.1 Hypothesis

For each ear one monaural model is run. Each ear perceps a certain loudness depending on the relation between the signal and the masker at each ear. Hypothesis: A test tone is audible, if the binaural loudness is above hearing threshold.

2.10.1.2 Stereo setup

To test the model for a binaural setup, the stereo setup described in section 2.2.1 was implemented in Matlab. The left and right ears' signals are determined by equations 10 to 13.

2.10.1.3 The panner

To achieve a constant overall loudness for all panning directions, a *Cosine² – Panner* has been chosen.

The angle of a sound event to arrive from lies between -30° and $+30^\circ$.

$$angle : -30^\circ \dots +30^\circ \quad (52)$$

The pan-factor has to reach values between 0 and 1, therefor the following equation is introduced.

$$pan = angle/60 + 0.5 \quad (53)$$

A pan-factor of 0 stands for hard left panning, 0.5 stands for center panning and 1 stands for hard right panning.

Let us assume that the masker is in center panning.

$$M = MR = ML \quad (54)$$

The left and right speaker signals SL and SR are derived from mono signal S. When the *cos* and *sin* functions take radiants as their arguments, SL and SR are given by the following equations.

$$SR = S * \cos((pan - 1) * \frac{\pi}{2})^2 \quad (55)$$

$$SL = S * \sin((1 - pan) * \frac{\pi}{2})^2 \quad (56)$$

The gain curves of the *Cosine² – Panner* are shown in figure 44.

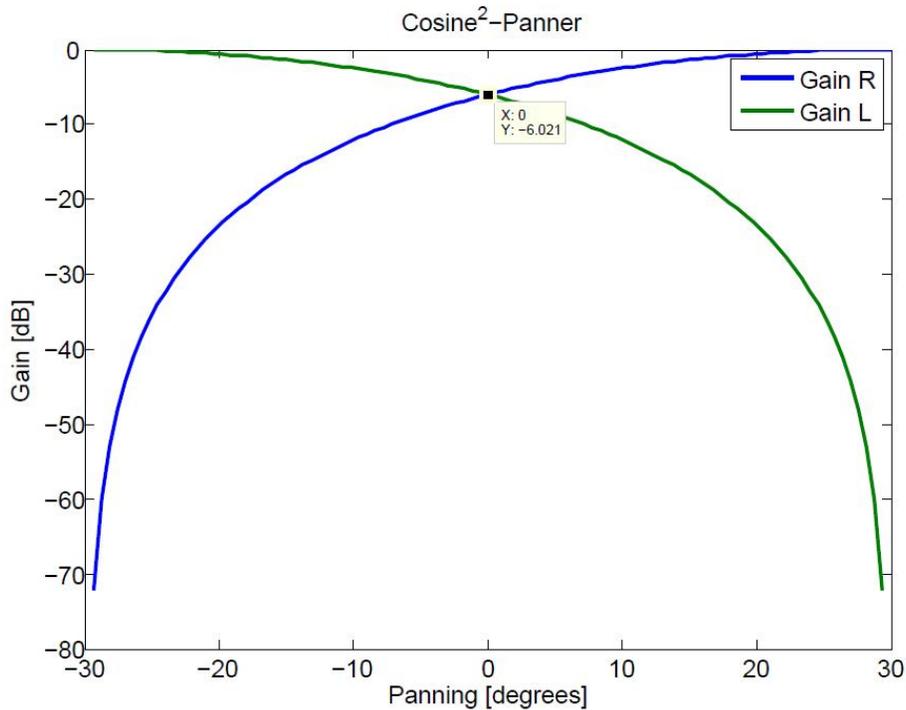


Figure 44: Gain factors of a *Cosine² – Panner*

Figure 45 shows the perceived loudness of a 1 kHz and 40 dB SPL sinusoid and a 26.4 dB SPL white noise, panned across the stereo triangle.

Let us now investigate the masking properties of a panned signal. The signal and masker are broadband, like in the empirical reference [15]. Figure 46 shows the relative level of a hard left panned masker, when the signal is at masking threshold.

2.10.1.4 Empirical data

Figure 47 is taken from [15] and shows the relative masker level for a constant speech intelligibility while varying the angle difference between masker and signal. Here the relative masker level differs from the predictions of the model. A reason for this can be that binaural unmasking (known as cocktail party effect) is not only caused by spectral masking. Beside that, the exact experimental conditions of the used literature are unknown. A way to objectively prove or

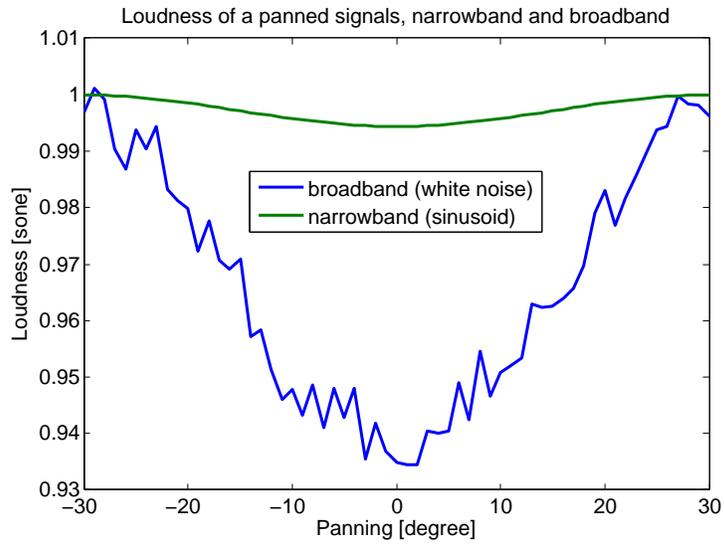


Figure 45: Model prediction: loudness of a panned sinusoid and a panned white noise

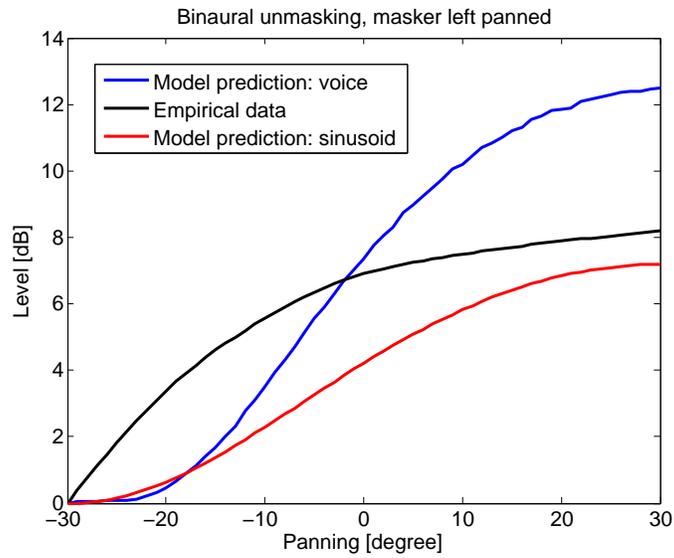


Figure 46: Model prediction versus empirical data [15]: binaural unmasking, masker (white noise) hard left panned, signal: sinusoid and voice

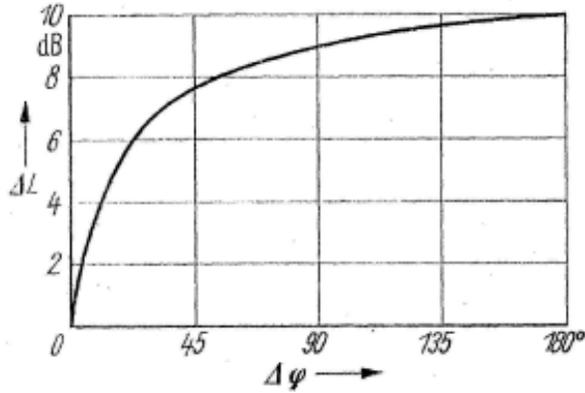


Figure 47: Empirical data for binaural unmasking taken from [15] (speech intelligibility), $\varphi = 0 - 60$ degrees corresponds to figure 46

disprove the hypothesis on binaural unmasking would be an objective experiment with enough qualified and trained listeners and known test conditions. This experiment is subject of further research.

2.11 Cross-analysis

When different instruments are mixed it is not targeting to decide one instrument to be a masker and another instrument to be masked. Rather each instrument works as a masker and is masked at the same time but in different spectral areas. To analyze those relations each instrument is once assumed to be masked by the sum of the rest of the instruments. In order to investigate these relations on the basis of an simple experiment, the number of instruments is defined to be two, so that once the excitation patterns for each instrument are known only the specific loudness calculation (see figure 1) has to be done again.

If the number of instruments were n , the roex-filters' input levels (see section 2.6) had to be calculated once for the sum of all instruments. Then $2 * n$ different excitation patterns had to be calculated, namely the excitation patterns for each instrument and for each supplementary sum of instruments (each instrument is masked by the sum of all other instruments).

Figure 48 shows the predicted partial loudness of a complex tone (sum of several sinusoids, see figure 48, first column) mixed with white noise. The lower

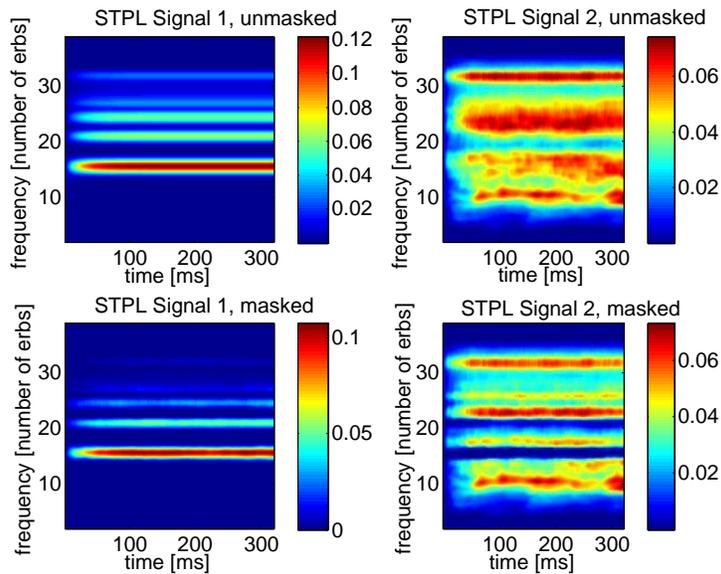


Figure 48: Cross analysis: specific loudness for white noise and a complex tone masking each other

subplots show the predictions for the same signals, but presented unmixed to the listener. In the upper two subplots one can see the suppression at both signals, interacting to each other. Dominant spectral parts of one signal mask the corresponding spectral areas of the other signal.

Figure 49 shows one time slice (1 ms) of figure 48.

Figure 50 shows one single time slice of partial loudness for two mixed white noises. Noteworthy is that areas of dominance are formed. At frequencies where one signal is dominant the other signal is suppressed and vice versa. Signal 1 nearly looks like a x-axis mirroring of signal 2. Areas where both signals are present are rare. Mixed signals seem to interact. Noise is very fluctuating in time and so areas of dominance change permanently. When stable sounds are mixed areas of dominance are stable.

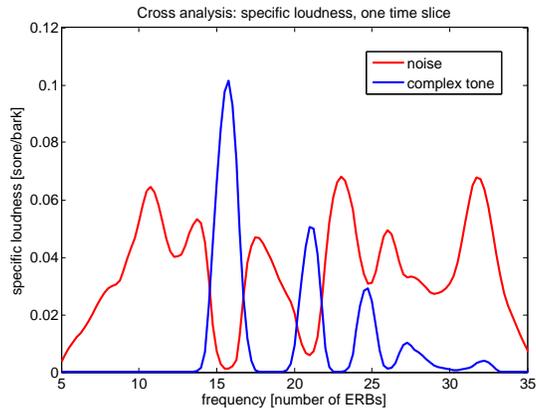


Figure 49: One time slice of figure 48

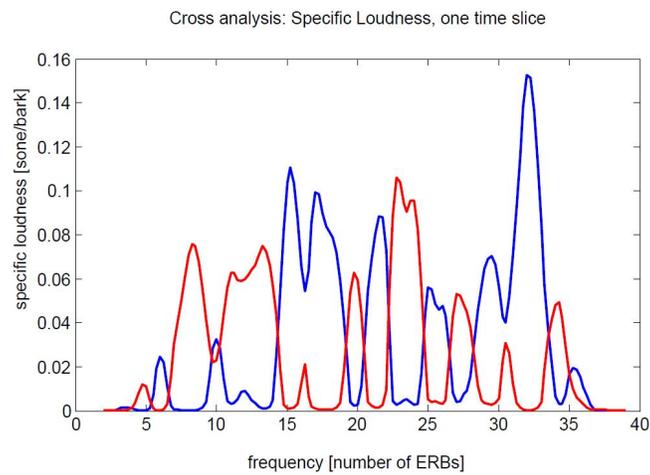


Figure 50: Cross analysis, one time slice: two white noises masking each other

2.12 Restrictions of the model

The model does not take into account the effect of loudness adaption that lets probands get used to high signal levels and make the perceived loudness smaller than predicted for longer periods of experiments [22]. Another restriction of the model is that the predicted masking threshold is not smeared by a temporal window to predict forward- and backward masking correctly [23]. That means short impulses have a greater effect on prior or subsequent events than predicted.

3 Experiment

3.1 Motivation

In order to find a measure for the quality of a mixed record the probability of correct identification of musical instruments (IP) was chosen to be the deciding factor. Some instruments of a mixdown can be discriminated very well, while other instruments affiliate to one unit. For automatic mixdown algorithms and for defining new user interfaces (see section 1) it is a goal to predict or control those inter-instrumental relations.

The described hearing model of section 2 is used to predict the loudness of masked (i.e. mixed) instruments.

3.1.0.5 Hypothesis

The identification probability of an instrument depends on how much of the original unmixed instrument is audible. The auditory model predicting specific loudness can also predict the IP of a mixed instrument.

3.1.1 Quotient of loudness for masked instruments

Let us define a new measure: the quotient of loudness for masked instruments (LQ) in percent gives the ratio between the total short-term loudness of an unmixed instrument and the total short-term loudness of the same instrument being mixed into a masking environment.

$$0\% \leq LQ = \frac{N_{stmasked}}{N_{stunmasked}} * 100\% \leq 100\% \quad (57)$$

where the short-term loudness variables of the masked and unmasked instrument $N_{stmasked}$ and $N_{stunmasked}$ are given by equation 32.

The aim of the following experiment is to find a relationship between LQ and IP.

Table 3: Experimental samplelist

Sample #	Instrument	Dynamics	Pitch
1	violin	piano	e2
2	violin	piano	a3
3	viola	piano	e1
4	viola	forte	a1
5	cello	forte	a
6	cello	piano	A
7	contrabass	forte	A
8	contrabass	piano	contra A
9	flute	forte	e1
10	flute	piano	a2
11	clarinette	forte	a
12	clarinette	forte	e1
13	oboe	forte	a1
14	oboe	forte	e1
15	bassoon	piano	A
16	bassoon	piano	E
17	trumpet	forte	a2
18	trumpet	piano	e2
19	horn	forte	e1
20	horn	forte	a
21	trombone	piano	a
22	trombone	forte	A
23	tuba	piano	e
24	tuba	forte	contra A

3.2 Testsignals

Recordings of 12 different instruments were thankfully provided by the author of [16]. The instruments are: violin, viola, cello, contrabass, flute, clarinet, oboe, bassoon, trumpet, horn, trombone and tuba. For each instrument two different samples were used. Therefore 24 different samples were tested. The samples are approximately 2 s long, the onsets of the instruments are original as recorded and the decays of the samples are faded out.

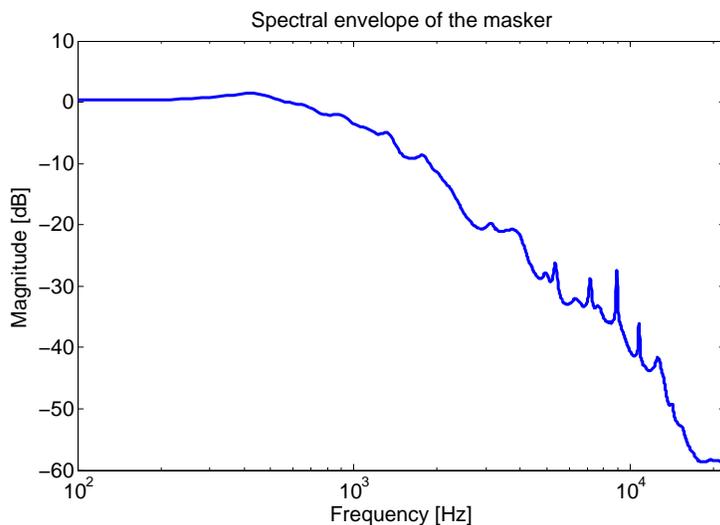


Figure 51: Spectral envelope of the masker

3.2.1 The spectral shape of the masker

An lpc^{10} algorithm was used to estimate the spectral envelopes of the samples. Matlab's $\text{lpc}()$ routine finds the coefficients of a 20th-order linear predictor (FIR filter) that predicts the current value of the real-valued time series x based on past samples. 20 coefficients were enough to form the spectral shape in an appropriate way. White noise was filtered with the average spectral shape of every sample. Advantage of this procedure is that the overall level of the masker could be kept low while masking was effective. See [17] for details of LPC. Figure 51 shows the plot of the masker's spectral envelope.

To create a masker that keeps LQ as constant as possible over time, colored noise was multiplied with the envelope of the instrument's signal.

The initial level of the noise determines LQ. The initial level was iteratively modified to reach the desired LQ. To keep data of the extracted tracks, all the analysis workspaces (i.e. excitations, spectral loudness, spectral centroid, spectral width, spectral deviation, short- and long term loudness ...) have been saved for later research on the empirical data. Figures 53, 54, 55 and 56 show some analysis plots for sample 3 (see table 3).

¹⁰ linear prediction coding

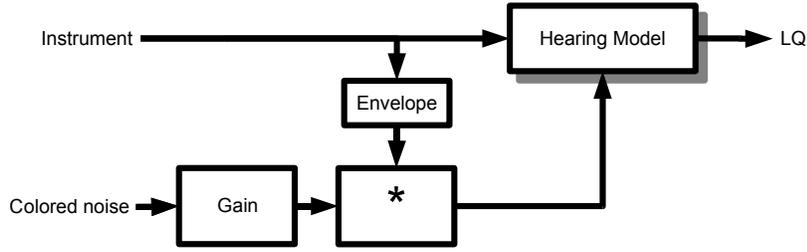


Figure 52: Masking the samples using different masker levels

The spectral centroid (figure 55) is given by

$$SC(t) = \frac{\sum_{f=0}^{fs/2} (N'_{sig}(f) * f)}{\sum_{f=0}^{fs/2} N'_{sig}(f)} \quad (58)$$

where $N'_{sig}(f)$ is specific loudness, f is the frequency and fs is the sampling frequency.

The spectral width (figure 56) is given by

$$SW(t) = f_l - f_u \quad (59)$$

where f_l is the lowest audible frequency of an instrumental sound and f_u is the highest audible frequency of an instrumental sound (audible means above 0.00537 [sone/bark]).

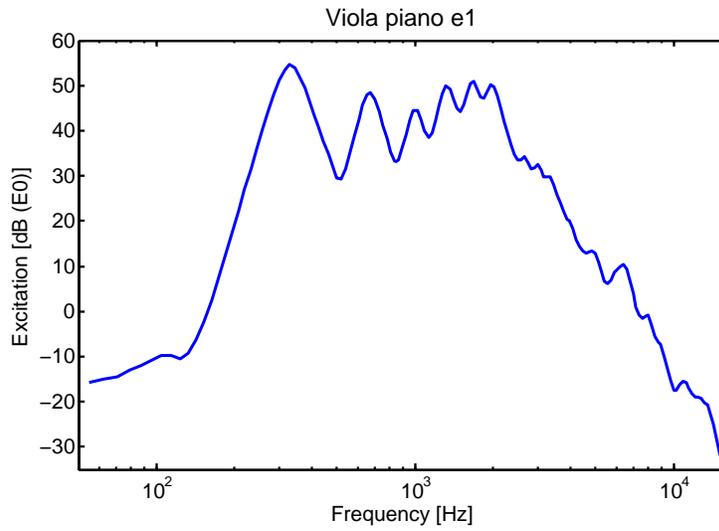


Figure 53: Excitation for an unmasked sample of a viola, playing a piano e1

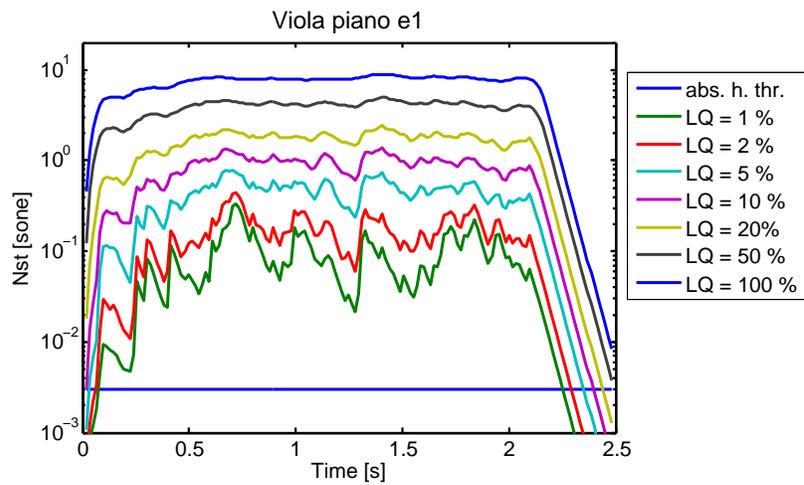


Figure 54: Short term loudness for an unmasked sample of a viola, playing a piano e1, different LQs

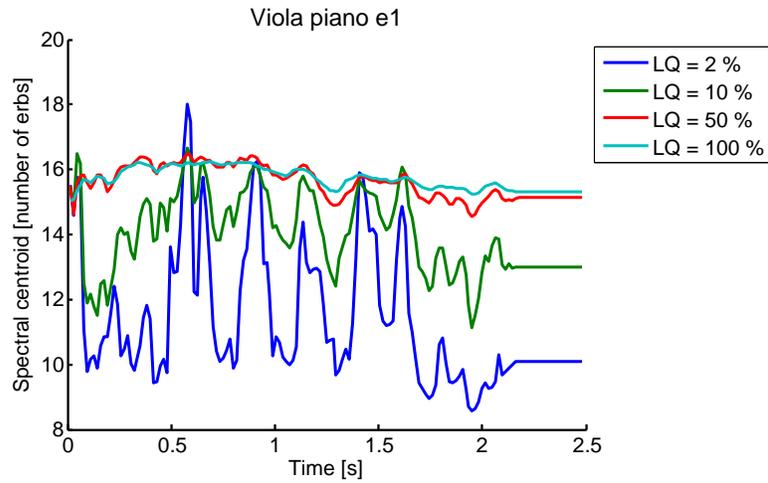


Figure 55: Spectral centroid for an unmasked sample of a viola, playing a piano e1, different LQs

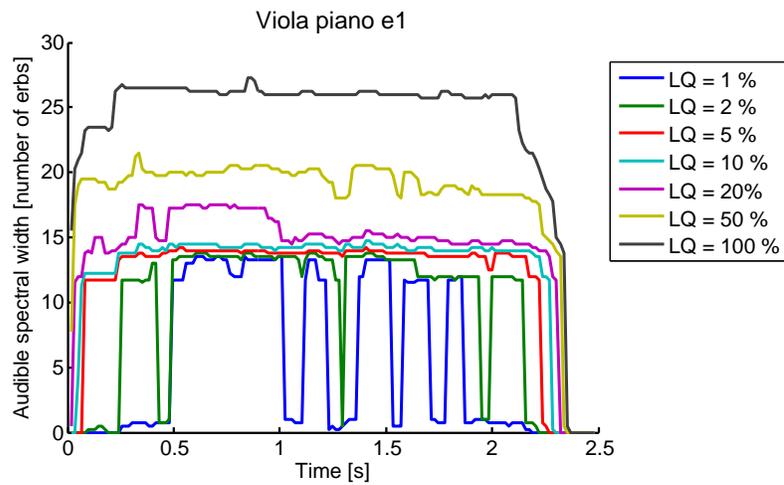


Figure 56: Spectral width for an unmasked sample of a viola, playing a piano e1, different LQs

3.3 Experimental design

In order to test the relation between LQ and IP, the probands were tested on trials of different levels of masking. Each trial contained one instrument (one sample of table 3) and the masker (section 3.2.1) at some known level. The probands were forced to identify what instrument is playing in each trial. The answers were given in a check box form, shown in the appendix.

Due to some preexperiments the function of LQ and IP was expected to be steeper for small LQs. Therefore the steps of LQ were chosen to be logarithmical. To change the probands' motivation for the better, some easier trials were added, i.e. trials at LQ = 30 % and 70 %. The tested LQs were 1 %, 2 %, 5 %, 10 %, 20 %, 30 %, 50 %, 70 % and 100 % (unmasked instrument). Due to the preexperiments the contrabass was also tested on LQ = 0.1 %, LQ = 0.2 % and LQ = 0.5 %, because the threshold was expected to be in that area. The loudest test signals were skipped at the test to ease the probands' ears. All probands were tested on 216 trials (24 samples * 9 levels of LQ + 3 contrabasstrials - 3 skipped). The trials were arranged in random order.

3.4 Report

3.4.1 Qualifying

At the analysis of the data two errors attracted attention.

The first error was when some proband was not able to identify some instrument at all, even for an unmasked scenario. That means that the answers of this proband for this instrument would only cause a variance to lower identification probabilities and bring no valid data.

The second error was when some proband was able to identify some instrument for all stages of masking, even when it was impossible in terms of timbre recognition. Those probands have memorized the temporal envelope of the masker for this instrument. Only timbre perception was subject of investigation and so answers of those two error types have been disqualified for further analysis.

For 10 samples more than 19 (50%) of the 38 probands are qualified. Those samples are in particular more meaningful than the others. Their sample numbers are 6, 7, 8, 10, 11, 12, 13, 18, 22 and 24 and will receive detailed analysis in section 3.4.3.1.

44 % of the probands' samples were excluded because of error 1 (proband was not able to identify some instrument at all) and 10 % were excluded because of error 2 (proband memorized the temporal envelope of the masker). The ratio between those two errors depends on the recorded raw material and is a trade off: More characteristic (i.e. containing temporal envelope cues) and longer samples

Table 4: Percentages of disqualification for samples, instruments and groups

Smpl. #	instr- ument	group	err- or 1	err- or 2	err- or 1 instr.	err- or 2 instr.	err- or 1 gr.	err- or 2 gr.
1	violin	str.	32%	42%	22%	61%	36%	21%
2	violin	str.	13%	79%				
3	viola	str.	53%	11%	46%	17%		
4	viola	str.	39%	24%				
5	cello	str.	63%	0%	50%	1%		
6	cello	str.	37%	3%				
7	c.bass	str.	29%	5%	24%	4%		
8	c.bass	str.	18%	3%				
9	flute	w.wind	84%	0%	49%	4%	42%	7%
10	flute	w.wind	13%	8%				
11	cl.tte	w.wind	29%	11%	25%	12%		
12	cl.tte	w.wind	21%	13%				
13	oboe	w.wind	39%	8%	46%	5%		
14	oboe	w.wind	53%	3%				
15	basson	w.wind	50%	3%	47%	5%		
16	basson	w.wind	45%	8%				
17	trumpet	br.	61%	3%	46%	1%	56%	3%
18	trumpet	br.	32%	0%				
19	horn	br.	71%	5%	72%	4%		
20	horn	br.	74%	3%				
21	t.bone	br.	79%	3%	61%	4%		
22	t.bone	br.	42%	5%				
23	tuba	br.	82%	0%	45%	1%		
24	tuba	br.	8%	3%				

of the instrument would minimize error 1 but maximize error 2. Smoothing out temporal envelope variances would minimize error 2 but maximize error 1.

Error 1 was primarily a problem for the horn (72 % disqualified) and the trombone (61 % disqualified). Error 2 was primarily a problem for the violin (61 % disqualified) and the viola (17 % disqualified).

Table 4 shows percentages of disqualification for samples, instruments and groups.

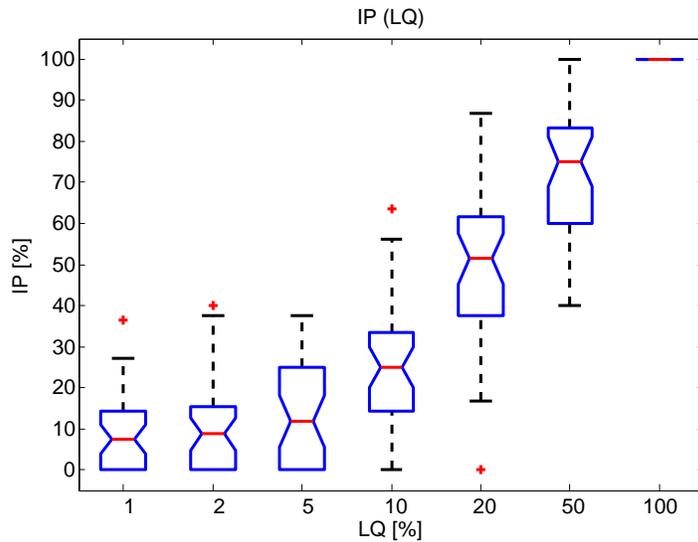


Figure 57: Anova boxplot for overall results, identification probability versus quotient of loudness

3.4.2 Overall results

3.4.2.1 AnOVA [18] and [19]

The analysis of variance (AnOVA) was executed in Matlab. It was proved that the identification probability IP [%] is related to the loudness quotient LQ [%].

To analyze the variance of the probands for each level of LQ, table 5 was fed in matlab's `anova1(X)`. Figure 57 shows the boxplot for this analysis. The red lines show estimators of the medians for each level of LQ. The boxes' notches show the confidence intervals in which the true medians lie with 95 % confidence. One can conclude, when the notches of two different boxes do not overlap, that the true medians are likely to be different.

The upper and lower ends of the boxes show the 25 % and 75 % - quartils (i.e. 25 % or 75 % of the values lie above). The whiskers (black lines) show 1.5 times the quartils or the maximum range of the values if there are no outliers (red crosses).

For LQ's of 10 % and larger results are extremely significant, for smaller LQ's the empirical data does not lead to significant results.

Table 5: Probability of correct instrument identification versus quotient of loudness for each proband, outliers are **bold**

Proband	LQ = 1 %	2 %	5 %	10 %	20 %	50 %	100 %
1	13.33	12.5	31.25	43.75	62.5	81.25	100.0
2	13.33	18.75	37.5	56.25	56.25	62.5	100.0
3	14.29	6.67	26.67	40.0	73.33	93.33	100.0
4	16.67	15.79	26.32	21.05	63.16	84.21	100.0
5	18.75	11.76	17.65	35.29	52.94	88.24	100.0
6	8.33	7.14	23.08	14.29	57.14	85.71	100.0
7	7.69	26.67	6.67	26.67	86.67	66.67	100.0
8	0	0	7.69	30.77	38.46	76.92	100.0
9	27.27	15.38	15.38	15.38	61.54	84.62	100.0
10	36.36	16.67	16.67	33.33	58.33	75.0	100.0
11	0	20.0	33.33	46.67	60.0	80.0	100.0
12	7.14	13.33	6.67	26.67	66.67	86.67	100.0
13	18.18	9.09	8.33	8.33	58.33	83.33	100.0
14	0	15.38	7.69	38.46	46.15	76.92	100.0
15	0	11.11	33.33	22.22	66.67	88.89	100.0
16	20.0	8.33	25.0	8.33	50.0	75.0	100.0
17	0	0	18.18	63.64	45.45	54.55	100.0
18	0	8.33	33.33	50.0	58.33	91.67	100.0
19	0	8.33	16.67	33.33	33.33	58.33	100.0
20	0	12.5	25.0	25.0	62.5	62.5	100.0
21	8.33	8.33	16.67	16.67	33.33	75.0	100.0
22	0	33.33	14.29	14.29	57.14	71.43	100.0
23	14.29	11.11	11.11	22.22	22.22	66.67	100.0
24	0	8.33	0	33.33	58.33	83.33	100.0
25	0	37.5	0	25.0	50.0	62.5	100.0
26	0	9.09	0	27.27	36.36	81.82	100.0
27	0	0	0	10.0	40.0	80.0	100.0
28	0	0	11.11	11.11	77.78	55.56	100.0
29	8.33	0	0	16.67	16.67	66.67	100.0
30	20.0	0	0	40.0	20.0	60.0	100.0
31	12.5	0	0	25.0	50.0	50.0	100.0
32	12.5	0	12.5	0	37.5	50.0	100.0
33	0	0	0	20.0	20.0	40.0	100.0
34	25.0	0	0	20.0	0	60.0	100.0
35	0	0	0	0	42.86	57.14	100.0
36	16.67	12.5	0	12.5	37.5	50.0	100.0
37	0	0	33.33	33.33	66.67	100.0	100.0
38	0	40.0	0	0	40.0	60.0	100.0

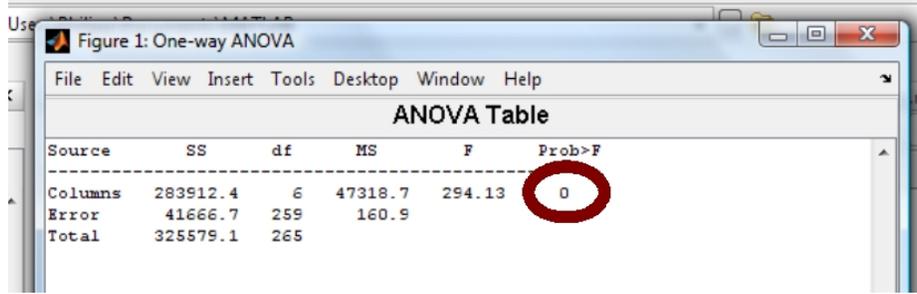


Figure 58: F-Test for overall results, results are highly significant

3.4.2.2 Modeling $IP(LQ)$

Matlab's function `regress.m` [18] was used to estimate a linear regression model for the observed results. Figure 59 shows the linear regression model for data of $LQ \geq 10\%$. The line is given by equation 60. It is not an exact line graphically because the steps 10, 20, 50 and 100 are not exactly equidistant on a logarithmic scale ($\log(10) = 1$, $\log(20) = 1.3$, $\log(50) = 1.7$, $\log(100) = 2$).

$$IP[\%] = \beta_0 + \log(LQ(\%)) * \beta_1 \quad (60)$$

where $\beta_0 = -46.68$ and $\beta_1 = 72.16$.

Figures 60, 61 and 62 show the linear regression models errors (residuals) for 3 different LQs. The 3 distributions thrown together give $mean = -0.7861\%$, $median = -0.4813\%$ and $\sigma = 16.0945\%$

Figure 63 shows the anova boxplot for comparing these 3 distributions. The probability that all 3 distributions have the same median is 25.05 %. Therefore the linear regression model is accepted to be valid.

For low LQs IP converges to guessing probability $1/n$. The relationship $IP(LQ)$ altogether is described by equation 61.

$$IP(LQ) = \begin{cases} LQ \geq 10^{\left(\frac{1/n - \beta_0}{\beta_1}\right)} & \dots \beta_0 + \log(LQ) * \beta_1 * \frac{a}{100\%} \\ LQ < 10^{\left(\frac{1/n - \beta_0}{\beta_1}\right)} & \dots \frac{1}{n} * 100\% \end{cases} \quad (61)$$

where n is the number of possible instruments, $a = IP(100\%)$ depending on the quality of the recordings and the training of the proband, $\beta_0 = -46.68$ and $\beta_1 = 72.16$.

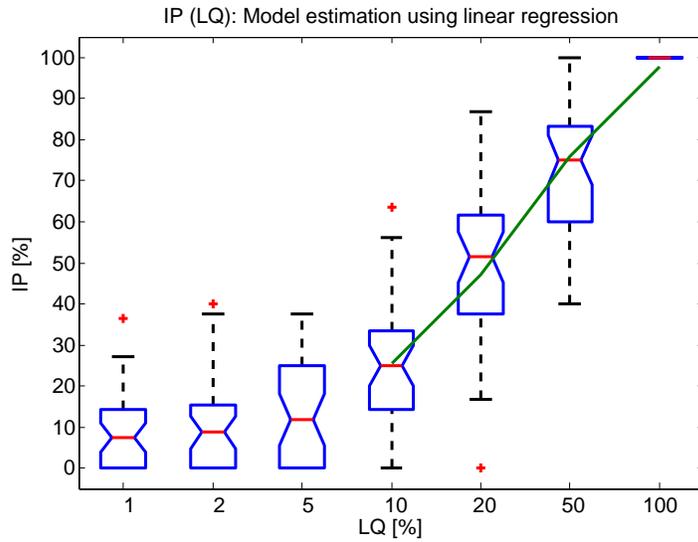


Figure 59: Data of $LQ \geq 10\%$ can be modeled by a linear regression model (green line)

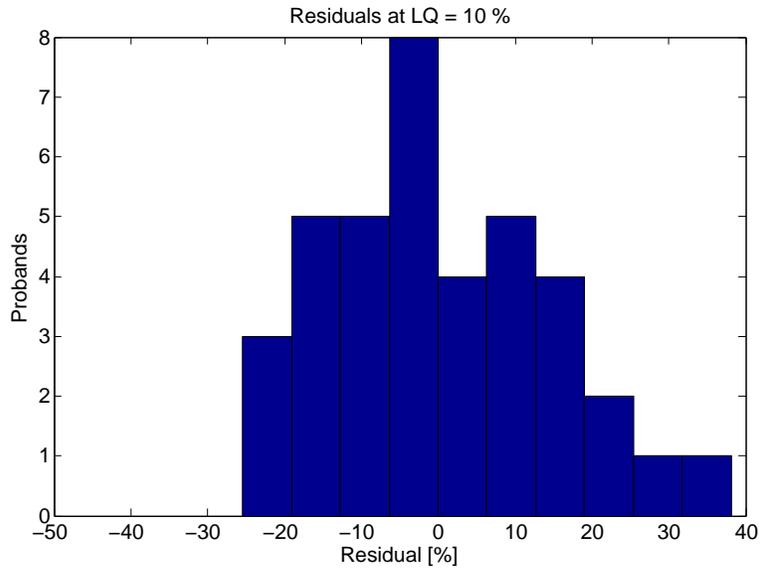


Figure 60: Residuals of the linear regression model and the empirical data at $LQ = 10\%$, $mean = -0.0387\%$, $median = -0.4813\%$, $\sigma = 15.1691\%$

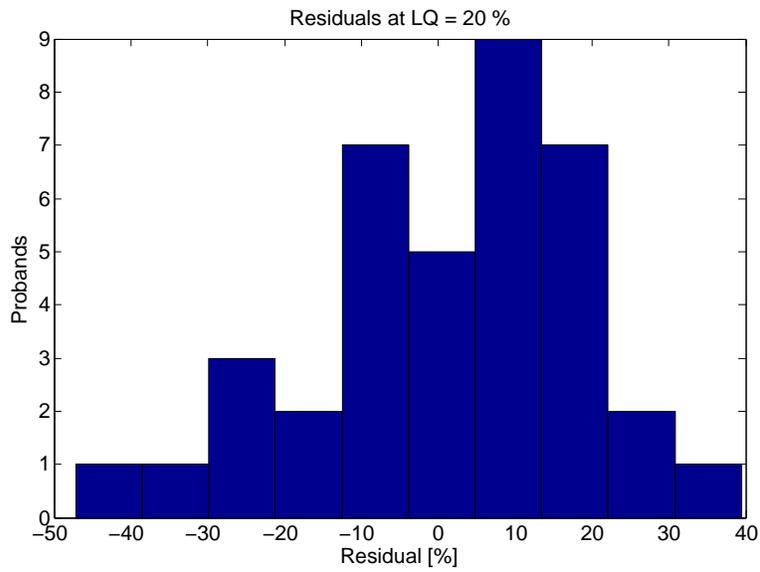


Figure 61: Residuals of the linear regression model and the empirical data at LQ = 20 %, $mean = 1.8517\%$, $median = 4.2662\%$, $\sigma = 18.2054\%$

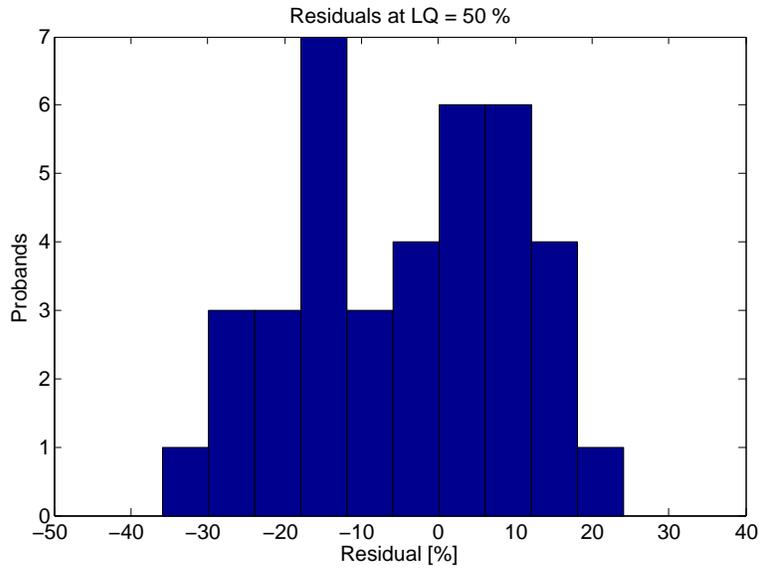


Figure 62: Residuals of the linear regression model and the empirical data at LQ = 50 %, $mean = -4.1712\%$, $median = -0.9194\%$, $\sigma = 14.4941\%$

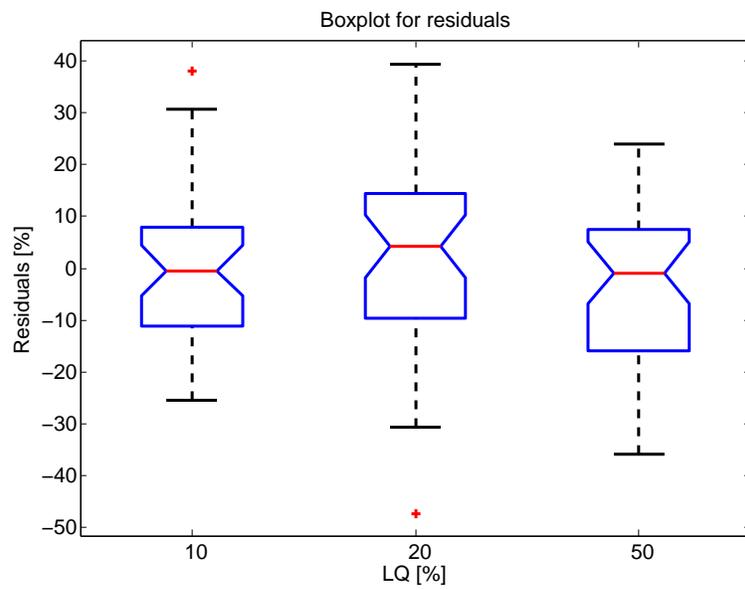


Figure 63: Boxplot of residuals, $F - Test : Prob > F : 25.05\%$

Table 6: Estimating the IP for single proband on single sample, 1 ... correct answer (correct instrument), 0 ... false answer

LQ	answer	averaged LQ	averaged answer
1 %	0	1.5 %	0.5
2 %	1	3.5 %	0.5
5 %	0	7.5 %	0
10 %	0	15 %	0.5
20 %	1	25 %	1
30 %	1	40 %	0.5
50 %	0	60 %	0.5
70 %	1	85 %	1
100 %	1		

3.4.3 Instrumental threshold determination

When analyzing single samples and single probands some treatment of probability estimation is necessary. To find the threshold where the $IP = 50\%$ a moving average procedure was used. Table 6 shows the answering and averaging of one proband on one sample. The threshold is found at the lowest LQ for which the averaged answer grows to 0.5 and not falls below that mark for any higher LQ again. Here it was estimated to be at $LQ = 15\%$.

To prove or unprove significant differences between the different samples another anova was executed. Figure 64 shows its boxplot.

A complete pair comparison of f-tests was executed. Its result is a matrix of significance (figure 65). Red cells show significant differences, green cells show significant similarities. Conclusions on this results are drawn in section 4.

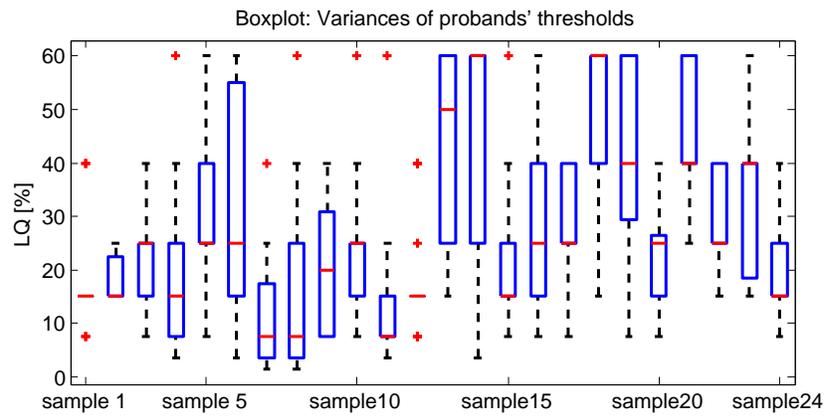


Figure 64: Boxplot: Variances of probands' identification thresholds for each sample

Sample	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
Sample 1		17%	7%	76%	2%	10%	9%	50%	29%	13%	4%	100%	0%	0%	38%	5%	0%	0%	0%	6%	0%	0%	2%	38%
Sample 2			44%	75%	15%	37%	7%	34%	76%	59%	3%	2%	2%	6%	94%	30%	11%	0%	6%	44%	1%	9%	15%	95%
Sample 3				9%	6%	31%	0%	0%	67%	61%	0%	0%	0%	0%	11%	31%	14%	0%	1%	56%	0%	12%	9%	7%
Sample 4					0%	4%	3%	23%	41%	14%	2%	62%	0%	0%	63%	2%	0%	0%	0%	0%	0%	0%	1%	55%
Sample 5						6%	0%	0%	12%	1%	0%	0%	8%	16%	0%	38%	43%	0%	26%	10%	10%	31%	100%	0%
Sample 6							0%	0%	35%	9%	0%	1%	2%	4%	4%	86%	93%	0%	12%	38%	6%	90%	64%	1%
Sample 7							3%	0%	0%	0%	81%	1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Sample 8								32%	9%	0%	40%	27%	0%	0%	6%	0%	0%	0%	0%	1%	0%	0%	0%	2%
Sample 9									9%	89%	1%	7%	1%	2%	51%	31%	14%	0%	3%	72%	1%	12%	15%	49%
Sample 10										89%	0%	0%	0%	0%	21%	9%	3%	0%	0%	70%	0%	2%	2%	15%
Sample 11									1%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Sample 12								40%	1%	0%	0%	0%	0%	0%	15%	0%	0%	0%	0%	0%	0%	0%	0%	15%
Sample 13								27%	7%	1%	0%	0%	0%	92%	0%	1%	1%	13%	79%	0%	0%	0%	16%	0%
Sample 14								0%	2%	0%	0%	0%	0%	0%	2%	2%	3%	16%	88%	1%	77%	1%	27%	0%
Sample 15								6%	51%	21%	0%	15%	0%	0%	2%	2%	0%	0%	0%	15%	0%	0%	0%	95%
Sample 16								0%	31%	9%	0%	0%	1%	2%	2%	86%	0%	6%	36%	2%	92%	48%	0%	0%
Sample 17								0%	14%	3%	0%	0%	1%	3%	0%	86%	0%	5%	16%	1%	90%	47%	0%	0%
Sample 18								0%	0%	0%	0%	0%	0%	13%	16%	0%	0%	0%	14%	0%	39%	0%	1%	0%
Sample 19								0%	3%	0%	0%	0%	0%	79%	88%	0%	5%	14%	1%	66%	2%	34%	0%	0%
Sample 20								0%	72%	70%	0%	0%	0%	1%	15%	36%	16%	0%	1%	0%	15%	12%	13%	0%
Sample 21								0%	1%	0%	0%	0%	0%	79%	77%	0%	2%	1%	39%	66%	0%	0%	14%	0%
Sample 22								0%	12%	2%	0%	0%	0%	0%	1%	92%	90%	0%	2%	15%	0%	36%	0%	0%
Sample 23								0%	15%	2%	0%	0%	0%	16%	27%	0%	48%	1%	34%	12%	14%	36%	0%	0%
Sample 24								2%	49%	15%	0%	15%	0%	0%	95%	0%	0%	0%	0%	13%	0%	0%	0%	0%
average of some group								average of same instr.	21,18%															
average of different group								average of different instr.	15,83%															

Figure 65: Matrix of significance: Variances of probands' identification thresholds for each sample. The thresholds of different samples differ significantly.

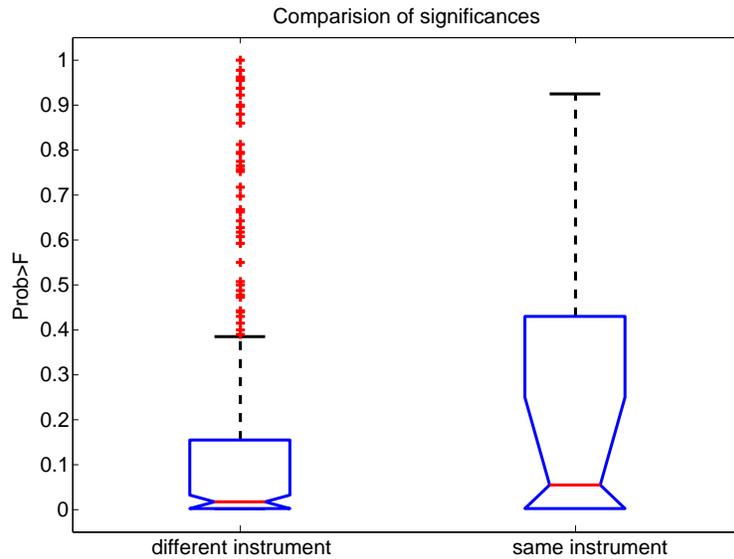


Figure 66: Samples of ident instruments seem to have more similar thresholds than samples of different instruments ($F - Test : Prob > F : 23.35\%$)

Figure 66 shows the boxplot of the values of the significance matrix, comparing ident instrument samples and different instrument samples. Two samples of ident instruments seem to have similar thresholds while samples of different instruments tend to have different thresholds.

Figure 67 shows the boxplot of the values of the significance matrix, comparing ident group samples and different group samples. Two samples of ident groups seem to have similar thresholds while samples of different groups tend to have different thresholds.

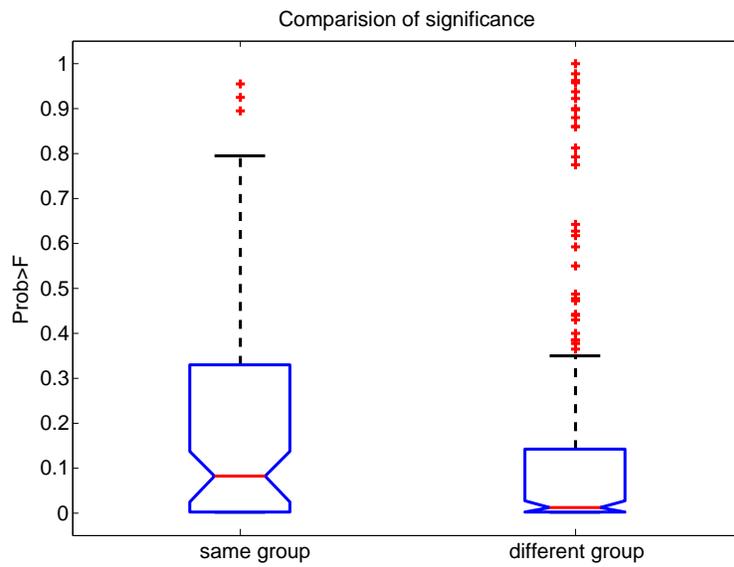


Figure 67: Samples of the same group seem to have more similar thresholds than samples of different groups ($F - Test : Prob > F : 8.16\%$)

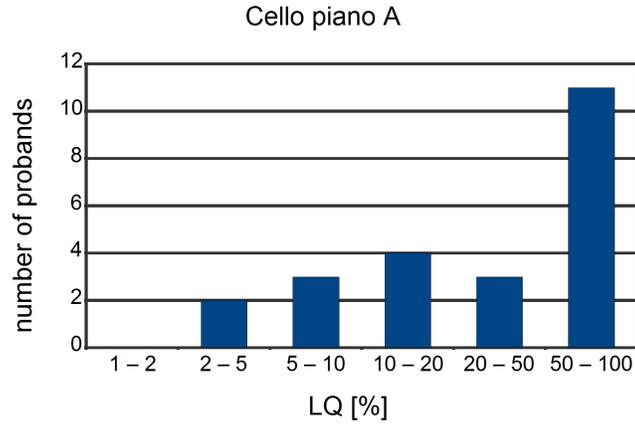


Figure 68: Identification threshold sample 6

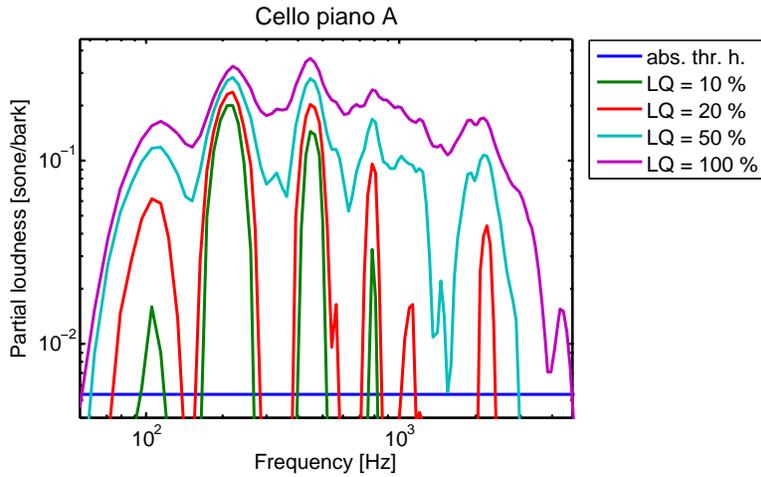


Figure 69: Partial loudness for sample 6

3.4.3.1 Detailed sample wise analysis

In this section the most meaningful samples (i.e. more than 50% of the probands qualified) are analysed in detail by the use of specific loudness patterns.

Sample 6: cello piano A

23 out of 38 probands qualified for this sample.

48 % of those probands had their threshold at $50\% < LQ < 100\%$ (atonal components between 3 kHz and 4.8 kHz). 17 % of those probands had their threshold at $10\% < LQ < 20\%$ (fifth, 11th and 22nd partial). $N_{stthr} = 2.60$ (see paragraph 3.4.3.2). $error_1 = 14$ and $error_2 = 1$.

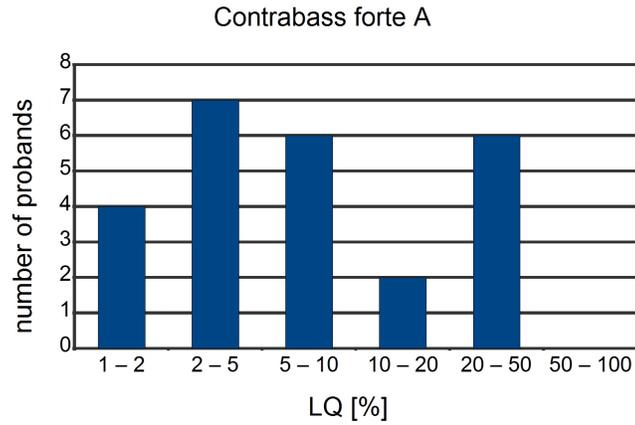


Figure 70: Identification threshold sample 7

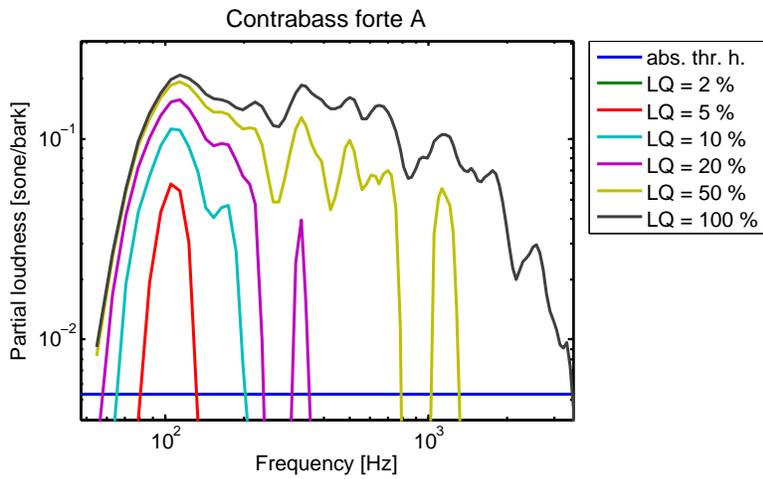


Figure 71: Partial loudness for sample 7

Sample 7: contrabass forte A

25 out of 38 probands qualified for this sample.

28 % of those probands had their threshold at $2\% < LQ < 5\%$ (fundamental). 24 % of those probands had their threshold at $5\% < LQ < 10\%$ (atonal components at 170 Hz). 24 % of those probands had their threshold at $20\% < LQ < 50\%$ (everything above the third partial). $N_{stthr} = 0.47$ (see paragraph 3.4.3.2). $error_1 = 11$ and $error_2 = 2$.

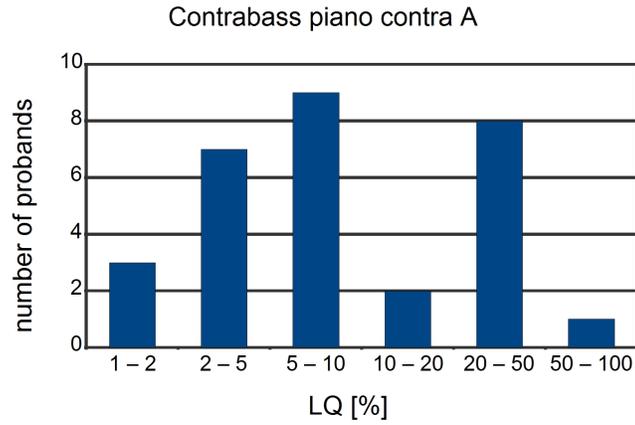


Figure 72: Identification threshold sample 8

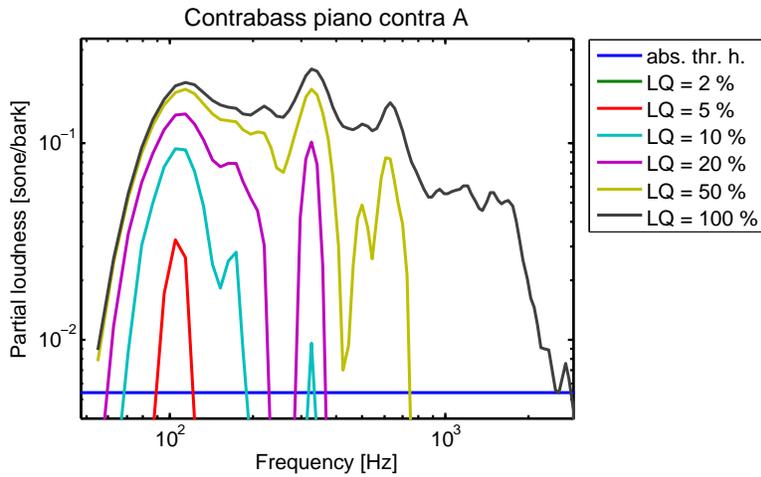


Figure 73: Partial loudness for sample 8

Sample 8: contrabass piano contra A

30 out of 38 probands qualified for this sample. 30 % of the probands had their threshold at $5\% < LQ < 10\%$ (third partial and atonal components at 170 Hz). 27 % of the probands had their threshold at $20\% < LQ < 50\%$ (fourth and fifth partial). 23 % of the probands had their threshold at $2\% < LQ < 5\%$ (fundamental). $N_{st_{thr}} = 0.54$ (see paragraph 3.4.3.2). $error_1 = 7$ and $error_2 = 1$.

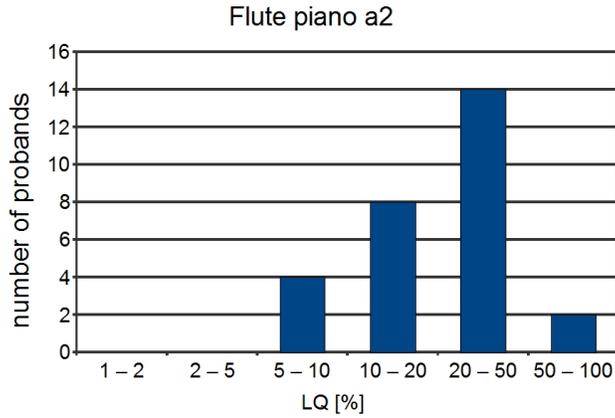


Figure 74: Identification threshold sample 10

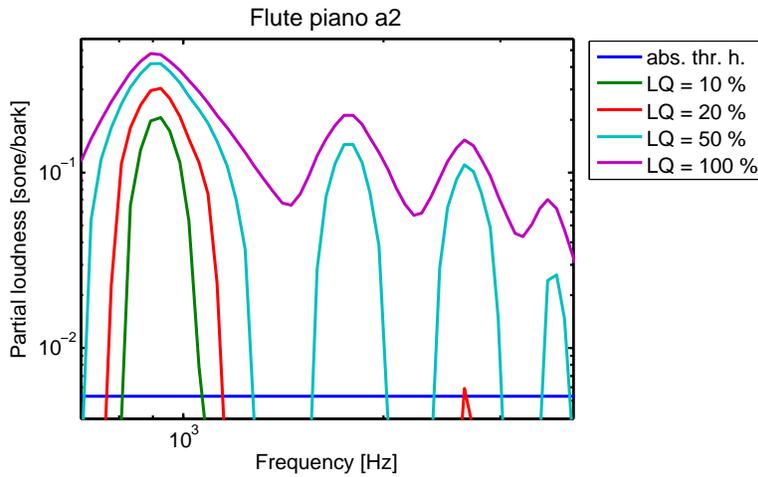


Figure 75: Partial loudness for sample 10

Sample 10: flute piano a2

30 out of 38 probands qualified for this sample. 50 % of the probands had their threshold at $20\% < LQ < 50\%$ (second and fourth partial). 29 % of the probands had their threshold at $10\% < LQ < 20\%$ (third partial). $N_{st_{thr}} = 1.14$ (see paragraph 3.4.3.2). $error_1 = 5$ and $error_2 = 3$.

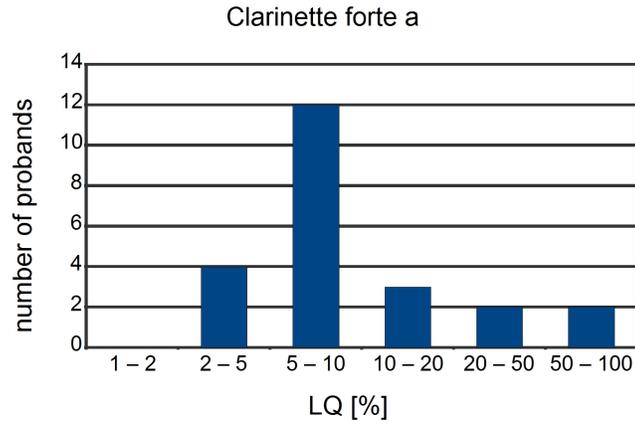


Figure 76: Identification threshold sample 11

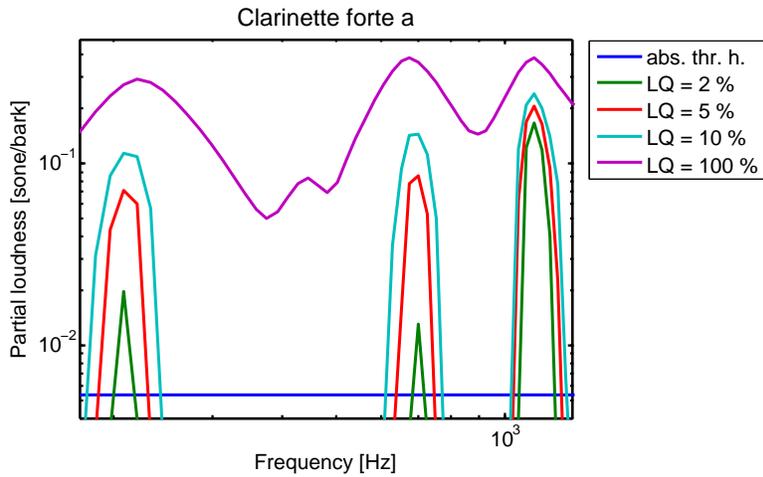


Figure 77: Partial loudness for sample 11

Sample 11: clarinette forte a

23 out of 38 probands qualified for this sample. 52 % of the probands had their threshold at $5\% < LQ < 10\%$ (just a reduction in overall loudness). 17 % of the probands had their threshold at $2\% < LQ < 5\%$ (just a reduction in overall loudness). $N_{st_{thr}} = 1.03$ (see paragraph 3.4.3.2). $error_1 = 11$ and $error_2 = 4$.

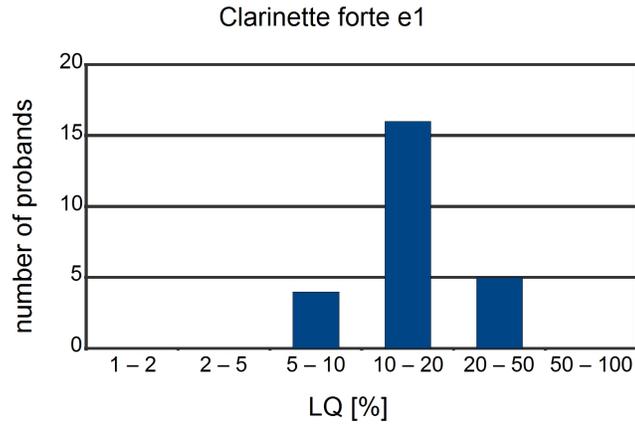


Figure 78: Identification threshold sample 12

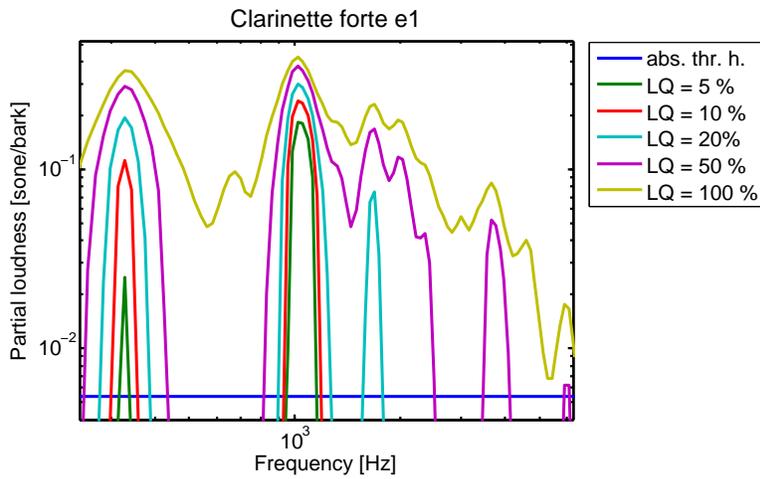


Figure 79: Partial loudness for sample 12

Sample 12: clarinete forte e1

25 out of 38 probands qualified for this sample. 64 % of the probands had their threshold at $10\% < LQ < 20\%$ (fifth partial). 20 % of the probands had their threshold at $20\% < LQ < 50\%$ (fourth, sixth, seventh, 11th and components around 6 kHz). 16 % of the probands had their threshold at $5\% < LQ < 10\%$ (just a reduction in overall loudness). $N_{stthr} = 1.03$ (see paragraph 3.4.3.2). $error_1 = 8$ and $error_2 = 5$.

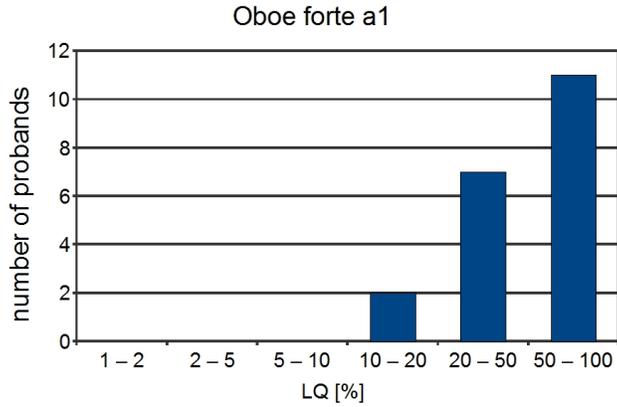


Figure 80: Identification threshold sample 13

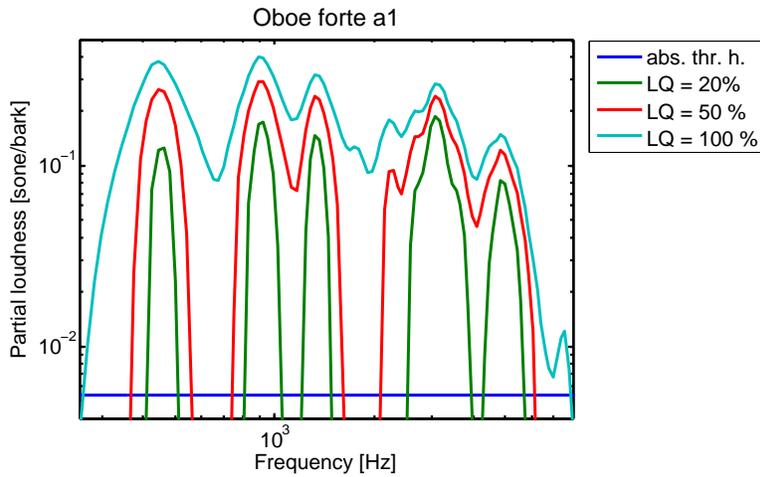


Figure 81: Partial loudness for sample 13

Sample 13: oboe forte a1

20 out of 38 probands qualified for this sample. 55 % of the probands had their threshold at $50\% < LQ < 100\%$ (fourth partial and components around 7.5 kHz). 35 % of the probands had their threshold at $20\% < LQ < 50\%$ (fifth partial). $N_{st_{thr}} = 3.85$ (see paragraph 3.4.3.2). $error_1 = 15$ and $error_2 = 3$.

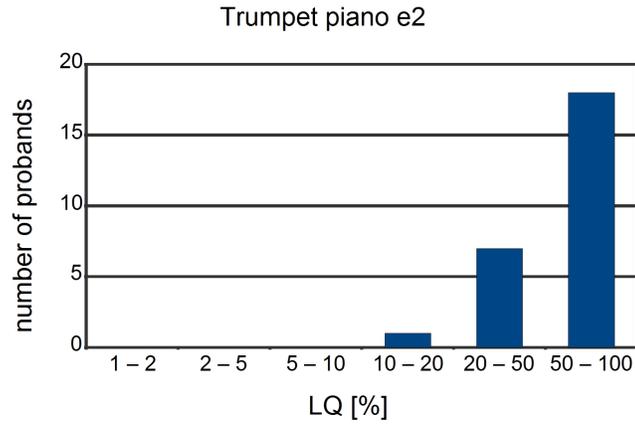


Figure 82: Identification threshold sample 18

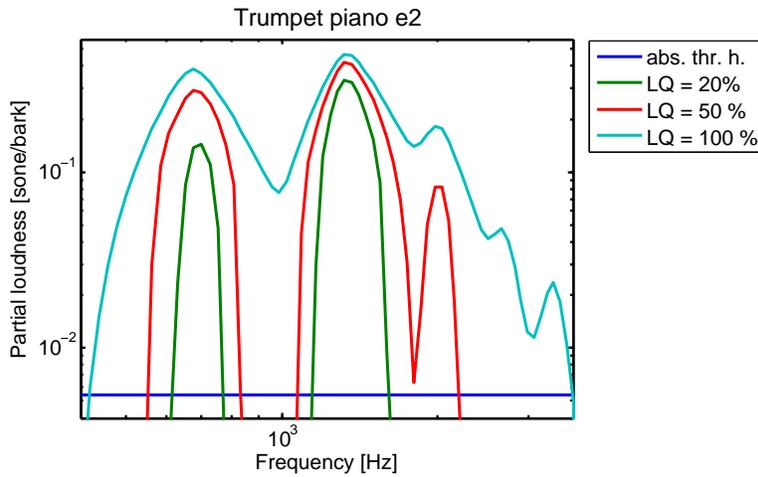


Figure 83: Partial loudness for sample 18

Sample 18: trumpet piano e2

26 out of 38 probands qualified for this sample. 69 % of the probands had their threshold at $50\% < LQ < 100\%$ (fourth and fifth partial). 27 % of the probands had their threshold at $20\% < LQ < 50\%$ (third partial). $N_{stthr} = 2.72$ (see paragraph 3.4.3.2). $error_1 = 12$ and $error_2 = 0$.

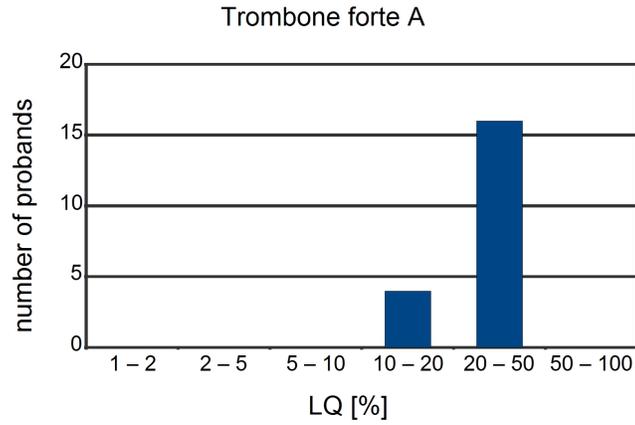


Figure 84: Identification threshold sample 22

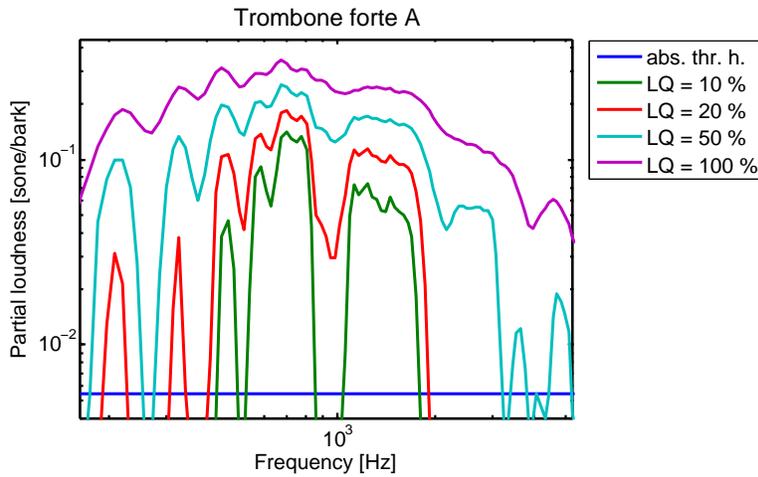


Figure 85: Partial loudness for sample 22

Sample 22: trombone forte A

20 out of 38 probands qualified for this sample. 80 % of the probands had their threshold at $20\% < LQ < 50\%$ (components between 2 and 3 kHz). 20 % of the probands had their threshold at $10\% < LQ < 20\%$ (second, third, eighth and ninth partial). $N_{sthr} = 2.36$ (see paragraph 3.4.3.2). $error_1 = 16$ and $error_2 = 2$.

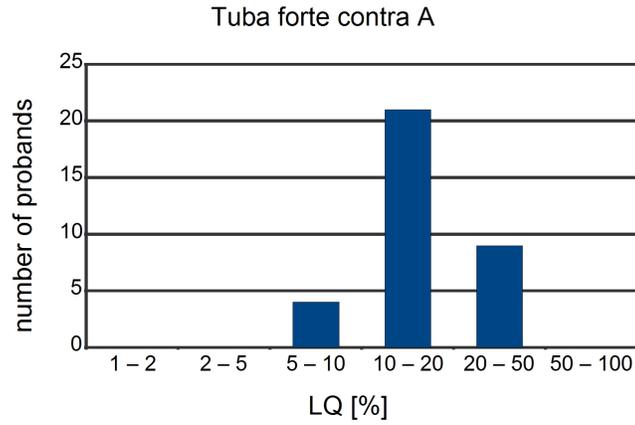


Figure 86: Identification threshold sample 24

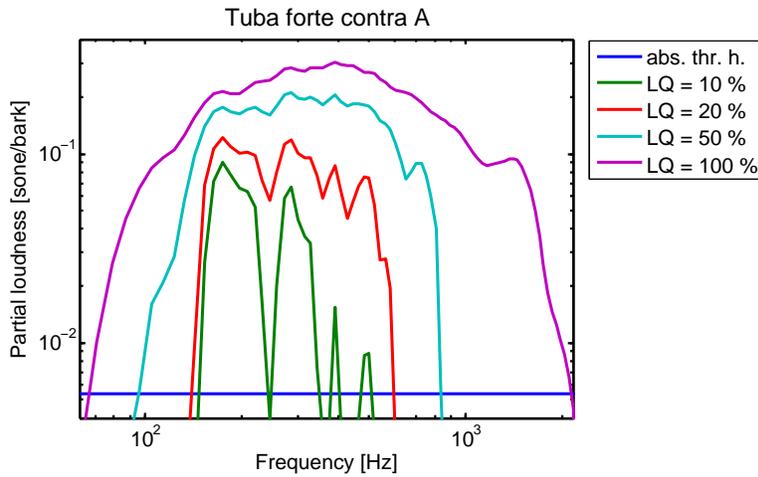


Figure 87: Partial loudness for sample 24

Sample 24: tuba forte contra A

34 out of 38 probands qualified for this sample. 62 % of the probands had their threshold at $10\% < LQ < 20\%$ (10th partial). 26 % of the probands had their threshold at $20\% < LQ < 50\%$ (second and 13th partial). $N_{st_{thr}} = 0.95$ (see paragraph 3.4.3.2). $error_1 = 3$ and $error_2 = 1$.

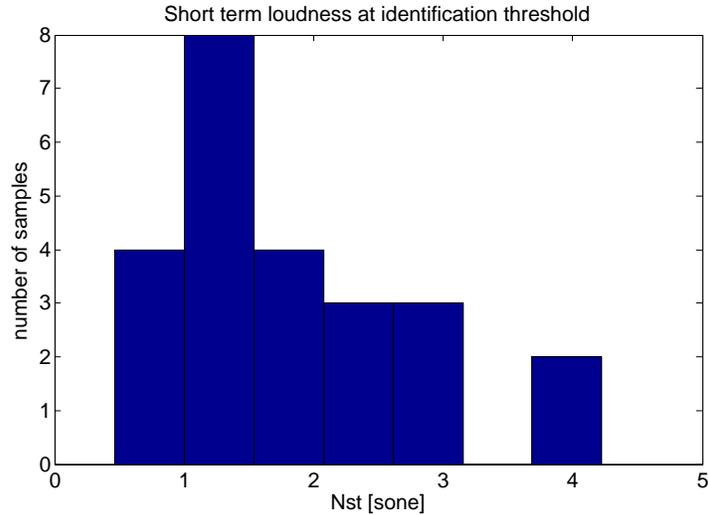


Figure 88: Overall short term loudness at identification threshold

3.4.3.2 Overall short term loudness at identification threshold

For each of the 24 samples the weighted short term loudness at the identification threshold was calculated using equation 62. Figure 88 shows the corresponding histogram.

$$N_{st_{thr}}(s) = \sum_{LQ=0.01}^1 N_{st}(s) * a(LQ)/Q \quad (62)$$

where s is the sample number, a is the number of probands who had their threshold at this LQ and Q is the number of probands who qualified for this sample s . It is noteworthy that the identification thresholds of different samples measured in absolute loudness instead of percents of LQ do match up more. This can be an indication to use absolute loudness in sone for further research also.

4 Conclusions and perspectives

Identification thresholds of some instruments (see for example figures 70 and 72) are subject to great personal variances. It seems that for those instruments different probands use different approaches to identify an instrument. Other instruments have less personal variance. Anyway it is not possible to find a threshold of LQ that fits for all instruments. Even some similar instruments (compare cello and contrabass) have markedly different thresholds. In general

one can say that two ident instruments playing different samples tend to have similar thresholds. Also samples of different instruments of an ident group do tend to have more similar thresholds than samples of instruments of different groups.

When looking at the general results (figures 57 and 88) one can see extremely significant tendencies. As a rule of thumb it is defined that the LQ of a mixed instrument should never fall below 0.1 or short-term loudness N_{stthr} should never fall below 1 some when the instrument should be perceptible separately. For soloing instruments a greater value should be obtained in order to bring more facets of timbre to the listener, and for background instruments (e.g. single voices of a choir) smaller values can be accepted.

The influence of temporal structures, panning and spatiality (i.e. reverb and echo) on masking needs more investigation. To define a precise instrumental identification model more research is planned also.

References

1. Moore B., Glasberg B., Baer T.: "A Model for the Prediction of Thresholds, Loudness, and Partial Loudness"; JAES Volume 45 Issue 4 pp. 224-240; April 1997
2. Moore B., Glasberg B.: "A Model of Loudness Applicable to Time-Varying Sounds"; JAES Volume 50 Issue 5 pp. 331-342; May 2002
3. Gonzales E. P., Reiss J. D.: "Automatic mixing: live downmixing stereo panner"; Proc. of the 10th int. conference on digital audio effects (DAFx-07), Bordeaux, France, September 10-15, 2007
4. Gonzales E. P., Reiss J. D.: "Improved control for selective minimization of masking using interchannel dependency effects"; Proc. of the 11th int. conference on digital audio effects (DAFx-08), Espoo, Finland, September 1-4, 2008
5. "EBU Tech. 3276 - 2nd edition"; May 1998
6. Parks T.W. and Burrus C.S.: "Digital Filter Design"; John Wiley & Sons, pp. 226, 1987.
7. Zwicker E., Fastl H.: "Psychoacoustics - Facts and models"; Springer, 1990
8. Painter T., Spanias A.: "Perceptual Coding for Digital Audio"; IEEE Volume 88, Issue 4, pp. 451-513, April 2000
9. Glasberg B. and Moore B.: "Derivation of auditory filter shapes from notched-noise data"; Hearing Research Volume 47, pp. 103-138, 1990
10. Cabrera D., Ferguson S. and Schubert E., "PsySound3: software for acoustical and psychoacoustical analysis of sound recordings"; Proceedings of the 13th International Conference on Auditory Display, Montreal Canada, June 26-29 2007, pp. 356-363.
11. Glasberg B. and Moore B.: "Development and Evaluation of a Model for Predicting the Audibility of Time-Varying Sounds in the Presence of Background Sounds"; JAES Volume 53, Issue 10, pp. 906-918, October 2005
12. Levitt H.: "Transformed Up-Down Methods in Psychoacoustics"; JASA Volume 49, Number 2, part 2, pp. 467-477, 1971
13. Smith D.: "MATLAB algorithm "Levitt 3-down,1-up - Method", University of Cambridge, Jan 2003, extended and enhanced by Pomberger H., IEM Graz, November 2008
14. Moore B. and Glasberg B.: "Modeling binaural loudness"; JASA Volume 121, Issue 3, March 2007

15. Fasold W.: "Taschenbuch Akustik, Teil 1"; VEB Verlag Technik Berlin, 1984
16. Jordan A.: "Akustische Instrumentenerkennung unter Berücksichtigung des Einschwingvorgangs, der Tonlage und der Dynamik"; Universität für Musik und darstellende Kunst Wien, Institut für Wiener Klangstil, Vienna, May 2007
17. Jackson L.B.: "Digital Filters and Signal Processing", Second Edition, Kluwer Academic Publishers, 1989. pp.255-257
18. The Mathworks. (2009) Statistics toolbox: Documentation. [Online]. Available: <http://www.mathworks.com/access/helpdesk/help/helpdesk.html> (15.9.2009)
19. Bech S., Zacharov N.: "Perceptual Audio Evaluation", John Wiley & Sons, 2006
20. Laback B.: "Skriptum zur Vorlesung Psychoakustik II Schwerpunkt: Experimentelle Audiologie", IEM Graz, 2006
21. <http://en.wikipedia.org/wiki/Timbre> (19.10.2009)
22. Carterette E.C.: "Loudness Adaptation for Bands of Noise", J. Acoust. Soc. Am. Volume 28, Issue 5, pp. 865-871 (September 1956)
23. Moore B.: "Characterization of Simultaneous, Forward and Backward Masking", 12th International Conference: The Perception of Reproduced Sound (June 1993), Paper Nr. 12-003

Voluntary details on the proband

Name or pseudonym: _____

Gender: Female Male

Age: _____

Profession: _____

Known loss of hearing _____

Instructions on the experiment

„Identification of masked instruments“

The experiment contains of 218 trials that are numbered from A03 to C20.

(Only for the alpha-group:

On Track A02 you will here unmasked samples of all instruments.)

The samples of the instruments will now be masked by noise of different levels. The order of playing the samples is random. Your task is to identify the right instrument for each trial and to check the corresponding box next to this instrument for each trial.

If you are not sure, try to guess the right group of instruments (strings, woodwind, brass). If you are absolutely unsure, simply guess.

Before every 10th track you will here a remark on the next tracknumber.

Overall duration: app. 40 minutes.

ANY QUESTIONS?

ANSWERS:

A03

strings:	<input type="checkbox"/> violine	<input type="checkbox"/> viola	<input type="checkbox"/> cello	<input type="checkbox"/> contrabass
woodwind:	<input type="checkbox"/> flute	<input type="checkbox"/> clarinette	<input type="checkbox"/> oboe	<input type="checkbox"/> bassoon
brass:	<input type="checkbox"/> trumpet	<input type="checkbox"/> horn	<input type="checkbox"/> trombone	<input type="checkbox"/> tuba

A04

strings:	<input type="checkbox"/> violine	<input type="checkbox"/> viola	<input type="checkbox"/> cello	<input type="checkbox"/> contrabass
woodwind:	<input type="checkbox"/> flute	<input type="checkbox"/> clarinette	<input type="checkbox"/> oboe	<input type="checkbox"/> bassoon
brass:	<input type="checkbox"/> trumpet	<input type="checkbox"/> horn	<input type="checkbox"/> trombone	<input type="checkbox"/> tuba

A05

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A06

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A07

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A08

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A09

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A10

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A11

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A12

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A13

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A14

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A15

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A16

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A17

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A18

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A19

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A20

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A21

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A22

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A23

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A24

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A25

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A26

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A27

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A28

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A29

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A30

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A31

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A32

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A33

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A34

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A35

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A36

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A37

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A38

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A39

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A40

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A41

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A42

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A43

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A44

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A45

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A46

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A47

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A48

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A49

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A50

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A51

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A52

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A53

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A54

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A55

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A56

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A57

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A58

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A59

strings: violine viola cello contrabass

woodwind: flute clarinette oboe bassoon

brass: trumpet horn trombone tuba

A60

strings: violine viola cello contrabass

woodwind: flute clarinette oboe bassoon

brass: trumpet horn trombone tuba

A61

strings: violine viola cello contrabass

woodwind: flute clarinette oboe bassoon

brass: trumpet horn trombone tuba

A62

strings: violine viola cello contrabass

woodwind: flute clarinette oboe bassoon

brass: trumpet horn trombone tuba

A63

strings: violine viola cello contrabass

woodwind: flute clarinette oboe bassoon

brass: trumpet horn trombone tuba

A64

strings: violine viola cello contrabass

woodwind: flute clarinette oboe bassoon

brass: trumpet horn trombone tuba

A65

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A66

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A67

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A68

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A69

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A70

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A71

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A72

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A73

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A74

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A75

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A76

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A77

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A78

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A79

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A80

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A81

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A82

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A83

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A84

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A85

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A86

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A87

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A88

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A89

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A90

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A91

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A92

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A93

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A94

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A95

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A96

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A97

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A98

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

A99

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B01

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B02

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B03

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B04

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B05

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B06

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B07

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B08

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B09

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B10

Streicher: Violine Viola Cello Kontrabass
Holzbläser: Flöte Klarinette Oboe Fagott
Blechbläser: Trompete Horn Posaune Tuba

B11

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B12

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B13

strings: violine viola cello contrabass

woodwind: flute clarinette oboe bassoon

brass: trumpet horn trombone tuba

B14

strings: violine viola cello contrabass

woodwind: flute clarinette oboe bassoon

brass: trumpet horn trombone tuba

B15

strings: violine viola cello contrabass

woodwind: flute clarinette oboe bassoon

brass: trumpet horn trombone tuba

B16

strings: violine viola cello contrabass

woodwind: flute clarinette oboe bassoon

brass: trumpet horn trombone tuba

B17

strings: violine viola cello contrabass

woodwind: flute clarinette oboe bassoon

brass: trumpet horn trombone tuba

B18

strings: violine viola cello contrabass

woodwind: flute clarinette oboe bassoon

brass: trumpet horn trombone tuba

B19

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B20

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B21

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B22

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B23

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B24

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B25

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B26

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B27

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B28

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B29

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B30

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B31

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B32

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B33

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B34

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B35

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B36

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B37

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B38

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B39

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B40

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B41

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B42

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B43

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B44

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B45

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B46

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B47

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B48

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B49

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B50

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B51

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B52

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B53

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B54

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B55

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B56

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B57

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B58

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B59

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B60

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B61

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B62

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B63

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B64

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B65

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B66

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B67

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B68

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B69

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B70

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B71

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B72

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B73

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B74

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B75

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B76

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B77

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B78

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B79

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B80

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B81

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B82

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B83

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B84

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B85

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B86

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B87

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B88

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B89

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B90

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B91

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B92

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B93

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B94

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B95

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B96

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

B97

strings: violine viola cello contrabass

woodwind: flute clarinette oboe bassoon

brass: trumpet horn trombone tuba

B98

strings: violine viola cello contrabass

woodwind: flute clarinette oboe bassoon

brass: trumpet horn trombone tuba

B99

strings: violine viola cello contrabass

woodwind: flute clarinette oboe bassoon

brass: trumpet horn trombone tuba

C01

strings: violine viola cello contrabass

woodwind: flute clarinette oboe bassoon

brass: trumpet horn trombone tuba

C02

strings: violine viola cello contrabass

woodwind: flute clarinette oboe bassoon

brass: trumpet horn trombone tuba

C03

strings: violine viola cello contrabass

woodwind: flute clarinette oboe bassoon

brass: trumpet horn trombone tuba

C04

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

C05

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

C06

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

C07

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

C08

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

C09

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

C10

Streicher: Violine Viola Cello Kontrabass
Holzbläser: Flöte Klarinette Oboe Fagott
Blechbläser: Trompete Horn Posaune Tuba

C11

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

C12

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

C13

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

C14

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

C15

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

C16

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

C17

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

C18

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

C19

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba

C20

strings: violine viola cello contrabass
woodwind: flute clarinette oboe bassoon
brass: trumpet horn trombone tuba