Laura Leucke

Estimating the Basilar Membrane Input-Output Function Using Fixed-Duration Masking Curves

Project's Thesis

Graz University of Technology

Institute for Electronic Music and Acoustics
Head: Dipl.-Ing. Dr.techn. Alois Sontacchi
Supervisor: MSc Ph.D. Georgios Marentakis

in coorporation with:
Acoustic Research Institut
Head: Doz. Dr. Peter Balasz
Supervisor: Dr. Thibaud Necciari

Graz, November 2015

Statutory Declaration

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

Graz, ____________________ ____________________

Date Signature

Eidesstattliche Erklärung[1]

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Graz, am ____________________ ____________________

Datum Unterschrift

[1]Beschluss der Curricula-Kommission für Bachelor-, Master- und Diplomstudien vom 10.11.2008; Genehmigung des Senates am 1.12.2008

Abstract

It is known very well that the auditory system, and especially the cochlea, reacts in a nonlinear way to incoming sounds. Nevertheless, some aspects regarding the nonlinearity of the basilar membrane (BM) are not fully understood yet.

Different models have been established estimating gain, compression and input-output function (I/O function) of the BM. As a matter of fact, the response of the BM evolves over the course of acoustic stimulation, an effect that is not taken into account in most models. One effect influencing BM response is the so-called medial olivocochlear reflex (MOCR). Based on experiments with animals, it is supposed that activation of the MOCR changes the BM I/O function. Precisely, the BM I/O function is shifted rightwards (i.e. towards high input levels) in comparison to the I/O function without influence of MOCR. As a result, a reduction in BM compression is expected. The results from previous studies with humans are inconclusive, though. In other words, the change in BM compression could not be confirmed.

In this study, this hypothesis has been investigated by a method called fixed-duration masking curves (FDMCs). As the method in its original form can not investigate the influence of MOCR on BM I/O function, the method has been adapted for this particular purpose. Two experiments were conducted. The results tend to confirm the change in BM compression with MOCR activation.

Contents

1 Theoretical Background **6**

1.1 The Peripheral Auditory System 6
1.2 The Medial Olivocochlear Reflex 9
1.3 Estimation of the I/O Function in Humans 11

2 Experiments **14**

2.1 Problem Description and Method 14
2.2 Listeners 15
2.3 Stimuli 16
2.4 Software Implementation 17
2.5 Procedure 17
2.6 XP1: Fixed-Duration Masking Curves with Cue Signal 19
2.7 XP2: Fixed-Duration Masking Curves without Cue Signal 23

3 Conclusion **27**

List of Figures

1 Cross section of the peripheral auditory system. 6

2 Production of a traveling wave along the BM for pure tone signals of different sound pressure levels (SPLs) [Ren, 2002]. 7

3 BM responses measured at a CF of 10 kHz for tones of different frequencies and levels in chinchillas [Oxenham and Bacon, 2003]. 7

4 BM I/O function for two different frequencies: a tone at CF (solid line) and a tone well below CF (dashed line). The inset illustrates the respective vibration patterns on the BM [Oxenham and Bacon, 2003]. 8

5 Schematic diagram of the broken-stick model [Yasin & Plack, 2003]. . . 9

6 Schematic diagram of MOC connections to the cochlea. 10

7 Schematic diagram of I/O function for activated and non-activated MOCR [Elizabeth A. Strickland, 2008]. 10

8 Schematic diagram of three different methods to estimate the BM I/O function in humans. 12

9 GOM method in the frequency domain. 12

10 FDMC method in the frequency (a) and time domain (b). 15

11 FDMCC method in the time domain. 19

12 Results of XP1 plotted as FDMCs. 20

13 Results of XP1 plotted as BM I/O functions (symbols). Broken-stick model fits (lines) were applied to the functions in the on- (red) and off-frequency precursor (black) conditions. 21

14 Result of XP2 for one listener plotted as FDMCs. 24

15 Results of XP2 plotted as I/O functions (symbols) for all listeners. Broken-stick model fits (lines) were applied to the functions in the on- (circles) and off-frequency precursor (squares) conditions. 25

1 Theoretical Background

1.1 The Peripheral Auditory System

1.1.1 Anatomy

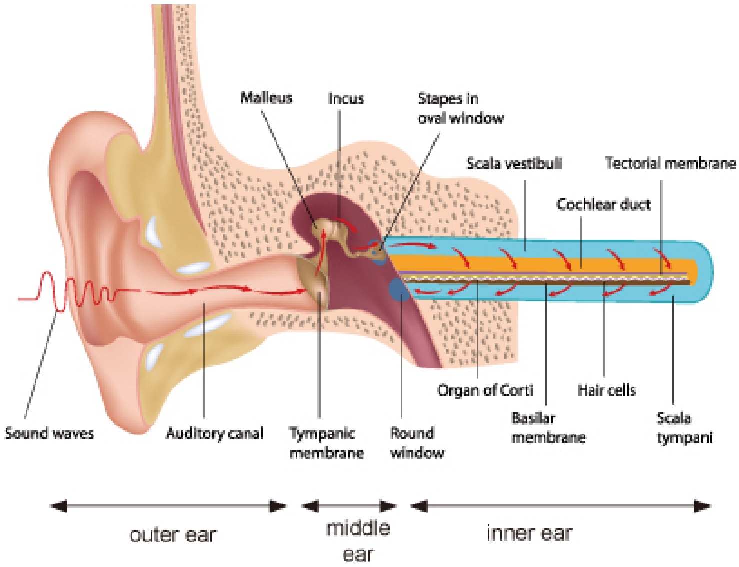


Figure 1: Cross section of the peripheral auditory system.

The peripheral auditory system is responsible for the hearing ability of humans who can perceive frequencies in the range from 20 Hz to 20 kHz. The auditory system is divided into three parts: the outer, middle and inner ear that are represented in Fig. 1. The outer ear, composed of the pinna and the auditory canal, acts as an amplifier for incoming sound waves. In addition, the pinna acts as a directional filter, which is important for the localization of acoustic sources.

The middle ear connects the eardrum with the oval window. Its main function is to realize an impedance matching between the air in the canal and the fluid in the cochlea. This is realized by the three ossicles: Malleus, Incus, and Stapes. Due to this impedance matching, humans can detect sound pressure variations from about $20\ \mu$Pa to about 20 Pa, the threshold of pain. This leads to a wide dynamic range of the auditory system. Furthermore, it can recover its full sensitivity within a fraction of a second from a very loud sound (Oxenham and Bacon, 2003).

The inner ear consists of the equilibrium organ (not shown in Fig. 1) and the cochlea. The main function of the cochlea is to transform mechanical waves in the fluid to electric signals for the neural processing of sounds in the central auditory system. This transformation can only succeed with the basilar membrane (BM). This is explained below.

1.1.2 The Basilar Membrane

When an acoustic signal reaches the eardrum, it produces a vibration pattern in the cochlea and a traveling wave spreads along the BM as shown in Figure 2.

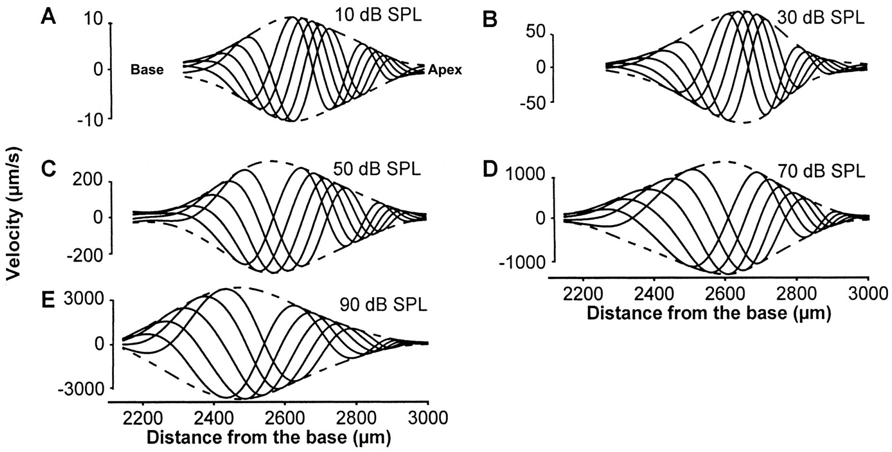


Figure 2: Production of a traveling wave along the BM for pure tone signals of different sound pressure levels (SPLs) [Ren, 2002].

The stiffness of the BM is different at every point, precisely it is high at the base and low at the apex. Consequently, the peak of the traveling wave will occur at a different position depending on the frequency of the tone. For high frequencies the peak is near the oval window, while for low frequencies it is near the apex. The frequency that produces the greatest response for a low-level sound at a given point on the BM is called the characteristic frequency (CF) of that point.

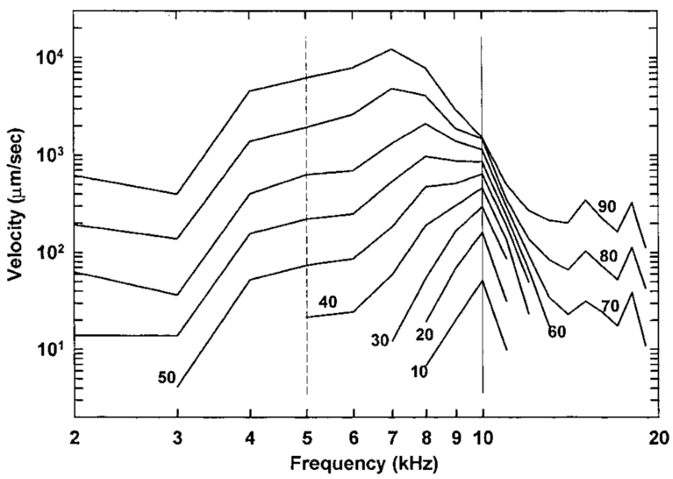


Figure 3: BM responses measured at a CF of 10 kHz for tones of different frequencies and levels in chinchillas [Oxenham and Bacon, 2003].

Figure 3 shows the BM response measured [2] at a CF of 10 kHz in chinchillas for tones of different frequencies and SPLs. It can be seen that the frequency selectivity decreases with increasing level: for low input levels the curves are very sharp whereas for high input levels they get much broader. It can also be seen that the change in response as a function of level is frequency dependent: for frequencies below CF the change in response is linear (a 10-dB increase in input level results in a 10-dB change in BM response), while for frequencies around CF the change in response is compressive (a 10-dB increase in input level results in less than 10-dB change in BM response). By plotting the BM response at a given CF as a function of the input SPL, one obtains the BM input-output function (I/O function) at that CF. Examples of BM I/O functions are shown in Fig. 4 based on the data plotted in Fig. 3. The solid line corresponds to the compressive I/O function at CF (10 kHz). The dashed line corresponds to the linear I/O function well below CF (5 kHz).

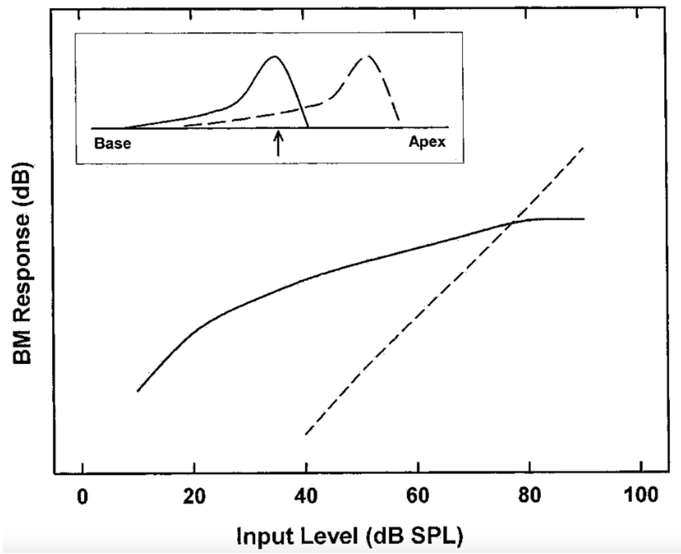


Figure 4: BM I/O function for two different frequencies: a tone at CF (solid line) and a tone well below CF (dashed line). The inset illustrates the respective vibration patterns on the BM [Oxenham and Bacon, 2003].

[2] The measurement was made with a laser velocimetry method.

1.1.3 The Broken-Stick Model

Yasin and Plack, 2003 proposed a model to describe the BM I/O function. The model is described by the following set of equations (see also Figure 5):

$$
\begin{aligned}
L_{out} &= L_{in} + \underbrace{(1-c)(BP_2 - BP_1)}_{G} && \text{if} \quad L_{in} \leq BP_1 \\
L_{out} &= cL_{in} + BP_1(1-c) + G && \text{if} \quad BP_1 \leq L_{in} \leq BP_2 \\
L_{out} &= L_{in} && \text{if} \quad BP_2 \leq L_{in}
\end{aligned}
\tag{1}
$$

L_{in} and L_{out} correspond to the input and output levels, respectively. G represents the gain of the I/O function and c is the slope of the function between BP_1 and BP_2. Because the model features three segments with discontinuities, it has been named the broken-stick model, which is shown in Fig. 5.

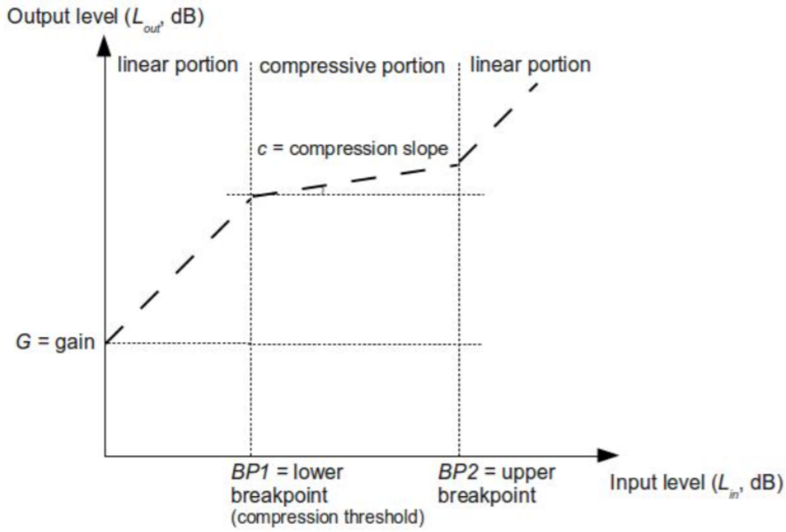


Figure 5: Schematic diagram of the broken-stick model [Yasin & Plack, 2003].

1.2 The Medial Olivocochlear Reflex

There is evidence that the BM I/O function as described in Sec. 1.1.2 is not static but evolves over the course of acoustic stimulation (Elizabeth A. Strickland, 2008). A potential reason for this non-stationarity is the activation of a feedback loop in the central auditory system, namely the medial olivocochlear (MOC) system. A simplified diagram of the MOC connections to the cochlea is shown in Fig. 6.

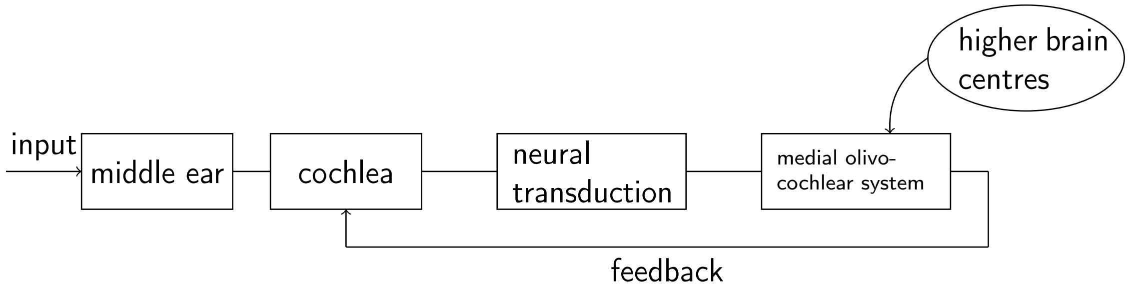


Figure 6: Schematic diagram of MOC connections to the cochlea.

The MOC system features both afferent (i.e. ascending from cochlea to brain) and efferent (descending from brain to cochlea) connections. The MOC system forms an acoustic reflex or feedback loop that, when activated through acoustic or electric stimulation, can reduce the BM response to a sound. Activation of the MOC reflex (MOCR) is controlled by higher brain centres. The main consequence of MOCR activation at the BM level is a reduction in gain G at low to mid levels (i.e. 20-50 dB SPL) according to Backus and J. J. Guinan, 2006. There is a time delay between the onset of the stimulus activating the MOCR and the onset of the gain reduction. In humans, this delay was estimated at about 25 ms (James et al., 2002; Backus and J. J. Guinan, 2006; Roverud and E. A. Strickland, 2010). Because the MOCR-induced gain reduction affects only the low levels (i.e. BP_2 is not affected), the BM I/O function as described by the broken-stick model is expected to change in the following way. First, there is an increase in BP_1. Second, because of the rightward shift of BP_1 and a constant BP_2, the slope of c must increase. Overall, the compressive region is reduced and the BM I/O function gets more linear. Figure 7 shows the expected changes in broken-stick model parameters with MOCR activation (compare dashed vs. solid lines).

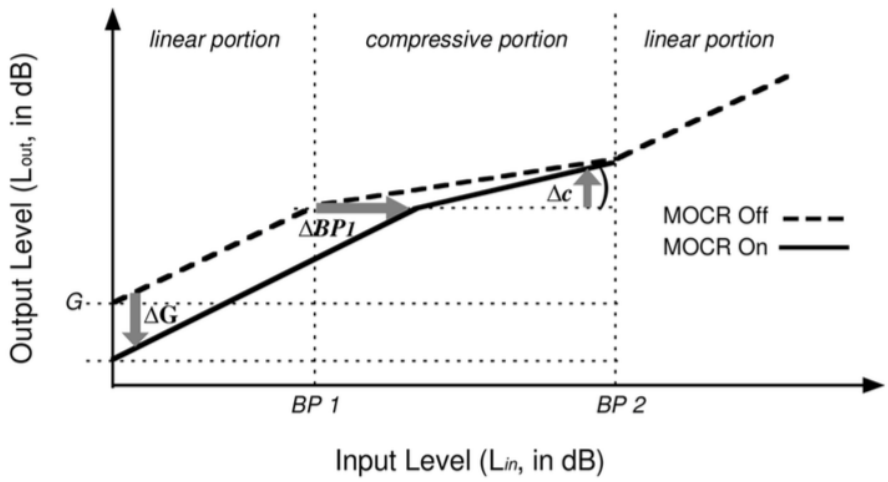


Figure 7: Schematic diagram of I/O function for activated and non-activated MOCR [Elizabeth A. Strickland, 2008].

While the changes in G and BP_1 have been consistently observed in both animals and humans (Cooper and Guinan, 2006;Jennings et al., 2009;Krull and E. A. Strickland, 2008), the change in c has only been observed in animals (Cooper and Guinan, 2006) and the results in humans are inconclusive (Jennings et al., 2009;Elizabeth A. Strickland, 2008;Roverud and E. A. Strickland, 2010). The aim of this study is to investigate the reduction in compression in humans when MOCR is activated.

1.3 Estimation of the I/O Function in Humans

In animals the I/O function can be easily measured with a laser detection method. Such a method cannot be used in humans because it is an intrusive method and nothing may be hurt in the cochlea. Therefore, an indirect method must be used such as auditory masking method.

Auditory masking refers to the situation where a sound (the "masker") reduces the audibility of another sound (the "target"). To measure masking, the detection threshold of the target is measured both in the presence (masked threshold) and absence of the masker (absolute threshold). The amount of masking in dB is thus given by the difference between the masked and the absolute threshold. Masking depends on the time, frequency, and level relationships between masker and target. For a detailed review on masking see Brian C. J. Moore, 2012, Chap. 3. BM functions can be estimated using simultaneous (masker and target are presented concurrently) or forward masking (the masker temporally precedes the target) Oxenham and Bacon, 2003. In this study we consider only forward masking.

Different forward masking methods have been established to estimate the BM I/O function in humans. We describe three common methods in the next sections. All three explained methods are depicted in the time domain in Figure 8. It is very essential to find a method that results in an estimate of the I/O function that is independent of efferent effects like the MOCR. First, it is necessary to separate the effects of the periphery and higher centres for better understanding of the auditory system, and second, reliable results can be used in cochlear models (Yasin, Drga, et al., 2013).

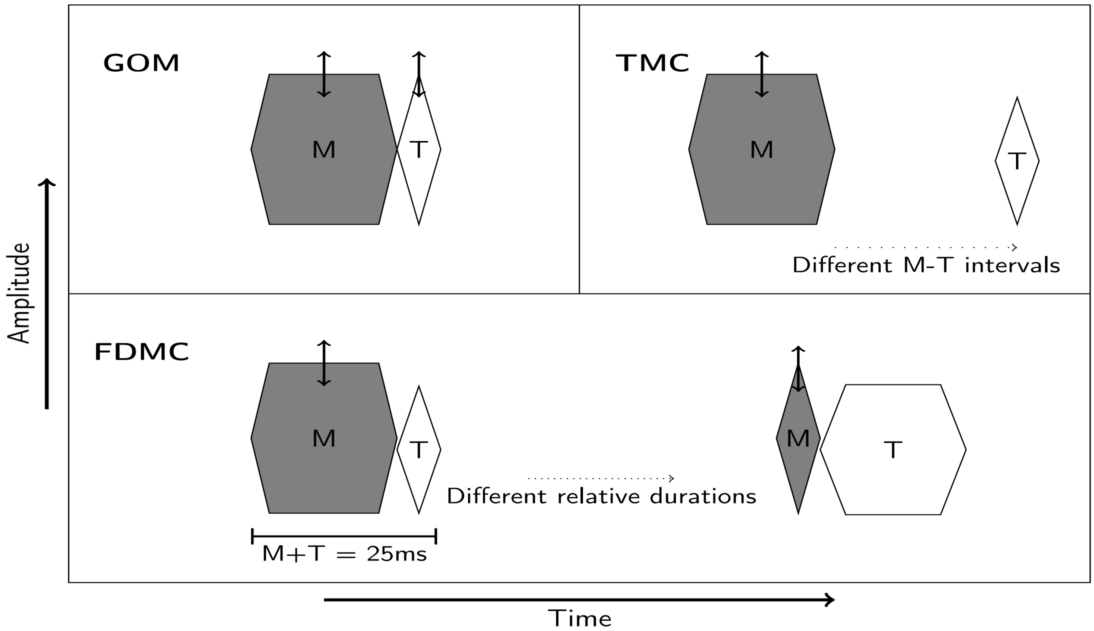


Figure 8: Schematic diagram of three different methods to estimate the BM I/O function in humans.

1.3.1 Growth of Masking Functions

The first method to explain is called "Growth of Masking" (GOM) method (Oxenham and Plack, 1997; Krull and E. A. Strickland, 2008; Roverud and E. A. Strickland, 2010; Elizabeth A. Strickland, 2008). Typically, a short tone burst with fixed frequency is used as target signal. The target frequency determines the point on the BM "where" the I/O function will be measured. A longer tone burst preceding the target is used as forward masker. For a given fixed target level, the masker level is adaptively varied until it masks the target. This procedure is repeated for a masker at the target CF (on-frequency masker) and a masker well below the target CF (off-frequency masker), and for several target levels. For better understanding this method is shown in the frequency domain in Fig. 9 (the target is plotted in blue and the different maskers are plotted in red).

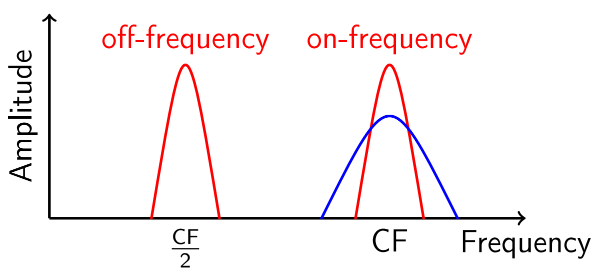


Figure 9: GOM method in the frequency domain.

By plotting the masker level at threshold as a function of the target level, one obtains a GOM function. Since the target level is assumed to be proportional to the input level and the masker level at threshold is assumed to be proportional to the amplitude of the BM response (output level), the shape of the GOM function can be used as an estimate of the BM I/O function (e.g. Oxenham and Plack, 1997). At the place of BM corresponding to the target frequency, the off-frequency masker is assumed to be processed linearly. In contrast, on-frequency masker is compressed at the target CF. The reasons have been explained in Sec. 1.1.2 (see Fig. 4).

To obtain a complete estimate of the function using GOM, high target levels are needed. This causes problems such as off-frequency listening that may bias the estimation of the function (Oxenham and Plack, 1997; Nelson et al., 2001). Thus, the GOM method is not ideal.

1.3.2 Temporal Masking Curves

The second method to explain is named "Temporal Masking Curves" (TMC) proposed by Nelson et al., 2001. A target signal with fixed frequency is presented at fixed low level. Then, the masker level at threshold is measured for various masker-to-target time intervals (M-T interval). Again, the procedure is repeated for on- and off-frequency forward maskers. The estimate of the BM I/O function is obtained by plotting the off-frequency masker level versus the on-frequency masker level at threshold paired by M-T interval (Nelson et al., 2001).

Because M-T intervals larger than 25 ms are needed to obtain a complete estimate of the function, the MOCR will be activated by the on-frequency masker. This leads to a gain reduction at the target CF at large M-T intervals. Thus, the result is not independent from efferent effects.

1.3.3 Fixed-Duration Masking Curves

A variant of the TMC has been proposed to avoid the implication of MOCR. The method is named "Fixed-Duration Masking Curves" (FDMC, Yasin, Drga, et al., 2013).

A target signal with fixed frequency and low level is presented. The masked threshold for an on- and off-frequency forward masker is measured. Here, the M-T interval is fixed at 0 ms, the masker plus target duration is fixed and the relative durations of masker and target are varied. By fixing the total M+T duration below or equal to 25 ms, efferent effects like MOCR are avoided. The I/O function estimate is obtained similarly to the TMC method, that is by plotting the off-frequency masker level versus the on-frequency masker level at threshold but paired by target duration instead as by M-T interval (see Yasin, Drga, et al., 2013 for details).

Overall, the FDMC method has the ability to provide an MOCR-free estimation of the I/O function if the total duration M+T is appropriately chosen, i.e. ≤ 25 ms. Therefore, we will opt for this method in our study.

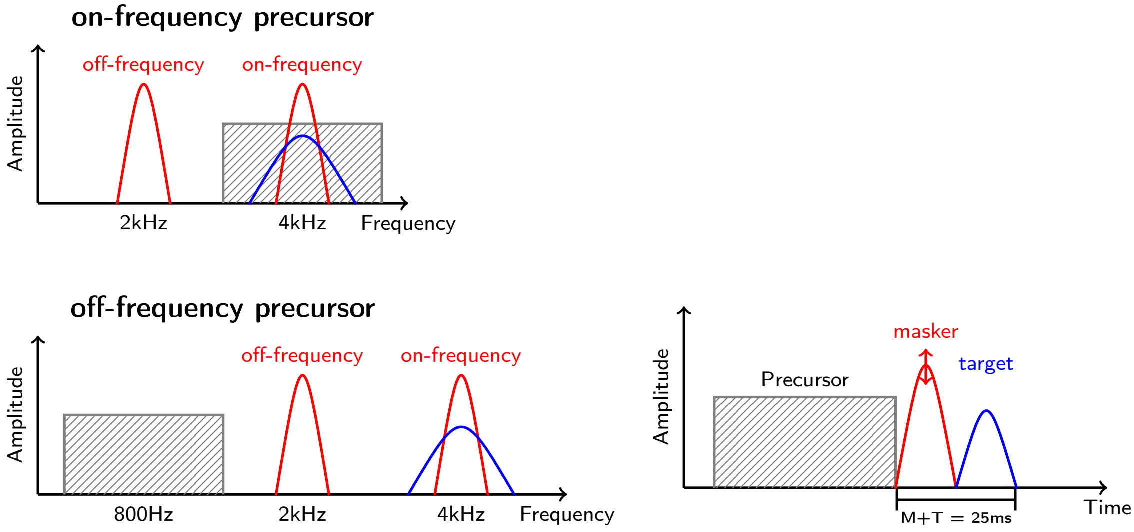
2 Experiments

2.1 Problem Description and Method

As stated in Sec. 1.2, the results of existing studies on I/O functions with activated MOCR in humans are inconclusive. The aim of this study is to assess the effect of MOCR on the BM response in humans.

The FDMC method can be adapted in different ways to measure two different characteristics of the BM I/O function. First, by setting M+T $\leq$ 25 ms the BM I/O function can be measured without the influence of efferents. Second, by adding a signal designed to trigger the MOCR, the BM I/O function can be measured in presence of efferents. The additional signal will be presented before the masker and target and is called a "precursor". To produce an efferent-induced gain reduction at the target CF, the spectral content of the precursor must be focused around the target CF (Lilaonitkul and Guinan Jr., 2009). The precursor level controls the strength of MOCR activation. The higher the precursor level, the greater the reduction in G (Guinan Jr., 2006). By comparing the results obtained in conditions with precursor and without precursor, one usually observes the effects of efferents on the BM I/O function. However, because the two conditions (with, without precursor) do not feature the same temporal sequence of stimuli (three signals in one case, two in the other case), a precursor signal is also used in the "reference" condition, i.e. without MOCR. To avoid the activation of efferents by the precursor in this condition, the precursor signal is moved far remote from the target CF. Therefore, there are two conditions: "Off-frequency precursor" and "on-frequency precursor".

The FDMC method is depicted respectively in the time and frequency domains in Figure 10.



(a) Frequency-Domain (b) Time-Domain

Figure 10: FDMC method in the frequency (a) and time domain (b).

Preliminary tests with a precursor resulted in unstable thresholds especially at short target durations. This instability might be due to the listeners having some difficulty to identify the target (Neff, 1985). In an attempt to fix this variability, a cue signal was added (see Sec. 2.6 for more details). The FDMC method with cue signal was the first experiment (XP1) to be run. Nevertheless, the experiment revealed some problems. The target level was not appropriately set and the cue signal did not reduce the variability. Therefore, a training period for all listeners was settled and the FDMC method was run without cue signal (XP2 - described in Sec. 2.7).

2.2 Listeners

Normal-hearing listeners with low absolute thresholds at 4kHz ($\leq$ 0 dB HL[3]) were selected for the experiments. This was done to allow for a large dynamic range of measurement. Indeed, participants with rather high thresholds could prevent us from measuring the low-level portion of the I/O function. In the first experiment, the FDMC method with cue signal (FDMCC), four listeners aged between 25 and 44 years participated. In the second experiment, ten listeners aged between 25 and 38 years participated. Three of them were very experienced in psychoacoustic testing.

All listeners were paid 7 Euro per hour for participation.

[3] dB HL ("hearing level"): dB relative to the threshold of the average normal-hearing population in the same age group.

2.3 Stimuli

Masker and target were sinusoids with 1.5 ms raised-cosine onset and offset ramps and a variable steady-state duration. The masker plus target duration was fixed at 25 ms. The shortest target was 3 ms (i.e. no steady-state portion). The longest target was 19 ms. The total masker duration (i.e. including the on- and offset ramps) thus varied between 22 and 6 ms in steps of 2 ms. The target frequency was set to 4 kHz. The on-frequency masker was set to 4 kHz and the off-frequency masker was set to 2 kHz.

The precursor was a band of narrowband noise with a bandwidth of two equivalent rectangular bandwidths (ERBs, Glasberg and B. C. J. Moore, 1990). Its duration was 150 ms. The on-frequency precursor had a center frequency of 4 kHz whereas the off-frequency precursor had a center frequency of 800 Hz.

For on- and off-frequency precursor conditions the target level was set to a fixed level. This level was different for the FDMC method with and without cue signal (see below). The masker start level was set to 20 dB SPL and varied according to the listener's response (see Sec. 2.5). The maximum masker level was set to 100 dB SPL. The precursor level was set to 40 dB SPL and remained constant for all conditions.

When present, the cue signal was a low-pass filtered noise with a cut-off frequency of 2 kHz. The cue signal was presented synchronously with the masker, therefore had the same duration as the masker. The cue signal level was set to 30 dB SPL (Neff, 1985). All stimulus characteristics are listed in Table 1.

	Target		**Masker**	**Precursor**	**Cue Signal**	
	FDMCC	FDMC			FDMCC	FDMC
Frequency [Hz]	4000	4000	2000;4000	800;4000	2000	-
Duration incl. ramps [ms]	3;5;7;9;11;13;15;17;19		22;20;18;16;14;12;10;8;6	150	22;20;18;16;14;12;10;8;6	
Level	10 dB SPL	10 dB SL	20 dB SPL (start level)	40 dB SPL	30 dB SPL	-
Bandwidth (proportion of center frequency)	-	-	-	0.25	0.99	-
onset-offset ramp [ms]	1.5	1.5	1.5	1.5	1.5	-

Table 1: Settings of all stimuli.

In principle the number of conditions amounts to 36 (9 durations x 2 maskers (on; off) x 2 precursors (on; off)). Nevertheless, long target durations produced very high masked thresholds, especially in the off-frequency masker condition. As the masker level was limited to 100 dB, some of the conditions could not be measured. Only 5 points were measured with the off-frequency masker. Overall, a total of 24 conditions remained for every listener during the experiment. All conditions were measured three times for each listener to get repeatable results.

2.4 Software Implementation

The whole regulation of the experiment was done with the software "ExpSuite". This software was developed at the acoustic research institute (ARI) in Vienna. With this software it is possible to generate different psychoacoustic experiment scenarios.

The software "ExpSuite" is based on visual basic and the whole procedure with all specifications (i.e. constants, variables) was implemented in this existing software. The sound files for each condition have been generated in MATLAB based on the input parameters provided by "ExpSuite". To supervise the experiment during the running process a table of constants and variables was defined. During the experiment all answers given by the listeners are stored in this table and are visible for the experimenter.

Listeners provided responses via a gamepad. The graphical interface already existed and had to be slightly adapted for the FDMC procedure.

The stimuli were digitally generated on a PC at a sampling rate of 48 kHz (24-bit resolution) and output via an external sound card E-MU 0202. Stimuli were presented to a Sennheiser HDA 200 headphone connected to a headphone amplifier (TDT HB6). Precursor, masker and target were presented to the left ear. The cue signal, when present, was presented to the right ear. The experiments were performed in a double-walled, sound-attenuated booth.

2.5 Procedure

Thresholds were measured using an adaptive three-interval, forced-choice task. The target was presented randomly in one of the three intervals in absolute threshold measurements, the other two intervals were silent. In masked threshold measurements, all three intervals contained the cue signal (in XP1 only), precursor and masker. The listeners had to indicate which interval sounded different from the other two by pressing one of three buttons of the gamepad. Each interval was visually signalled on the graphical interface, with a between-interval gap of 300 ms. Response feedback was provided after each trial by visually highlighting the interval containing the target. In addition, the correctness of the response was indicated by inserting either "correct" or "incorrect" on the screen.

In the adaptive procedure, the masker level was varied using the three-up one-down rule that estimates the 79.4% point on the psychometric function (Levitt, 1971). The masker level started at a sufficiently low level so that the target be easily audible. The initial step size was 10 dB and was halved after the second and fourth reversal. A run was terminated after 12 reversals, and the threshold was calculated by averaging the masker level over the last eight reversals. Runs with a standard deviation larger than 6 dB were dropped and re-run.

In a preliminary experiment the detection threshold of the 3ms target was measured in quiet (absolute threshold) and in presence of the on-frequency precursor (a 22 ms silent gap replaced the masker). In this setting the target level instead of the masker level varied adaptively. This step was necessary to set the appropriate SPL of the target in the main experiment.

All conditions were presented in a random order for each listener. Listeners were trained at least five hours on selected conditions. Some listeners needed up to eight hours of training until their thresholds became stable.

2.6 XP1: Fixed-Duration Masking Curves with Cue Signal

This experiment was programmed and conducted in August-September 2014.

2.6.1 Design

Figure 11 reminds the schematic diagram of the experiment setting in the time domain. For the detailed settings of the stimuli for the FDMCC experiment see Table 1.

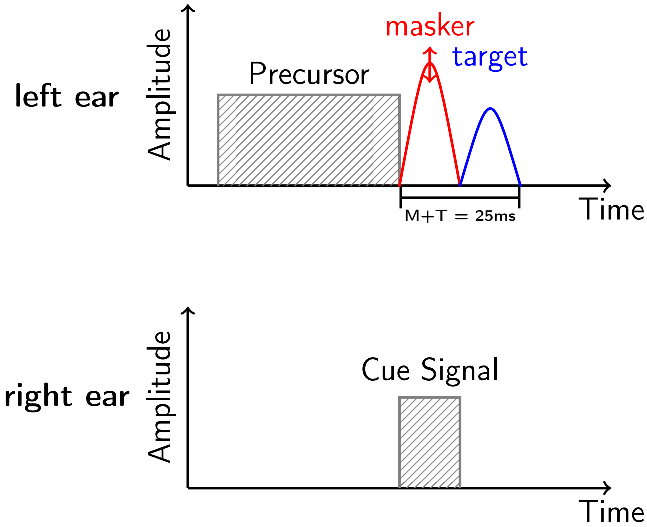


Figure 11: FDMCC method in the time domain.

In this experiment, the target level was set to 10 dB above the masked threshold of the 3 ms target measured in presence of the cue signal and the on-frequency precursor. This corresponded to 45-59 dB SPL across listeners. This target SPL was used in all conditions. Noteworthy, we verified that the cue signal presented to the contra-lateral ear did not mask the target.

2.6.2 Results and Discussion

The results for on- (left column) and off-frequency precursor (right column) for four listeners are shown as FDMCs in Fig. 12.

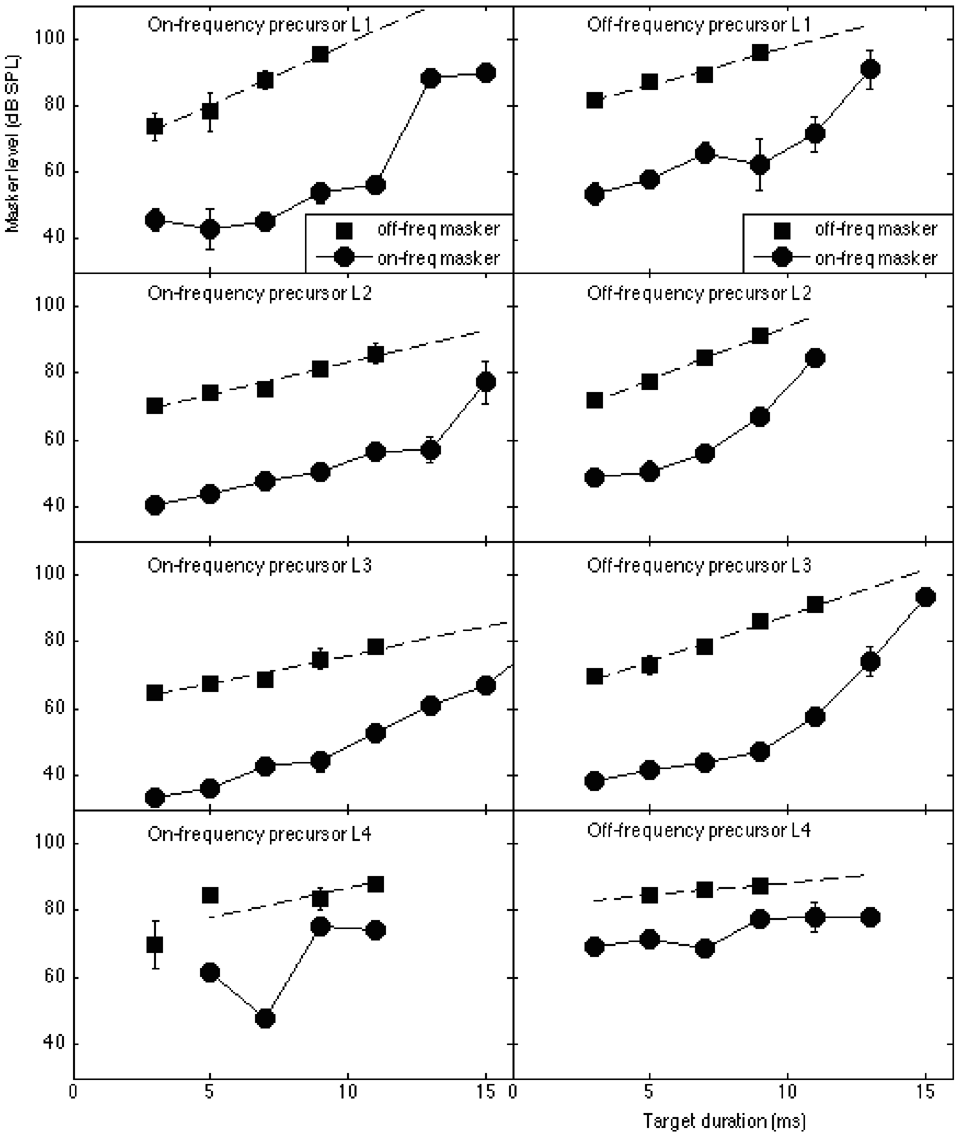


Figure 12: Results of XP1 plotted as FDMCs.

The dashed line for the off-frequency masker condition corresponds to the linear regression of the measured thresholds. The linear fit is necessary to derive the BM I/O function for all target durations. One can see that the results for the off-frequency masker are linear, as expected.

The standard deviation for each condition is illustrated by the error bars at each measured threshold. Although error bars are generally small, one can see that the variability did not vanish totally. In particular, the results for L1 reveal some variability for target durations ≥ 9 ms. One potential reason might be that, in those conditions, the cue signal is not clearly audible because the masker level gets very high. This was verified in a follow-up experiment (not shown).

The BM I/O functions derived from the FDMCs of the four listeners are shown in Figure 13. All functions were fitted with the broken-stick model described in Eq. (1) (straight lines in Fig. 13).

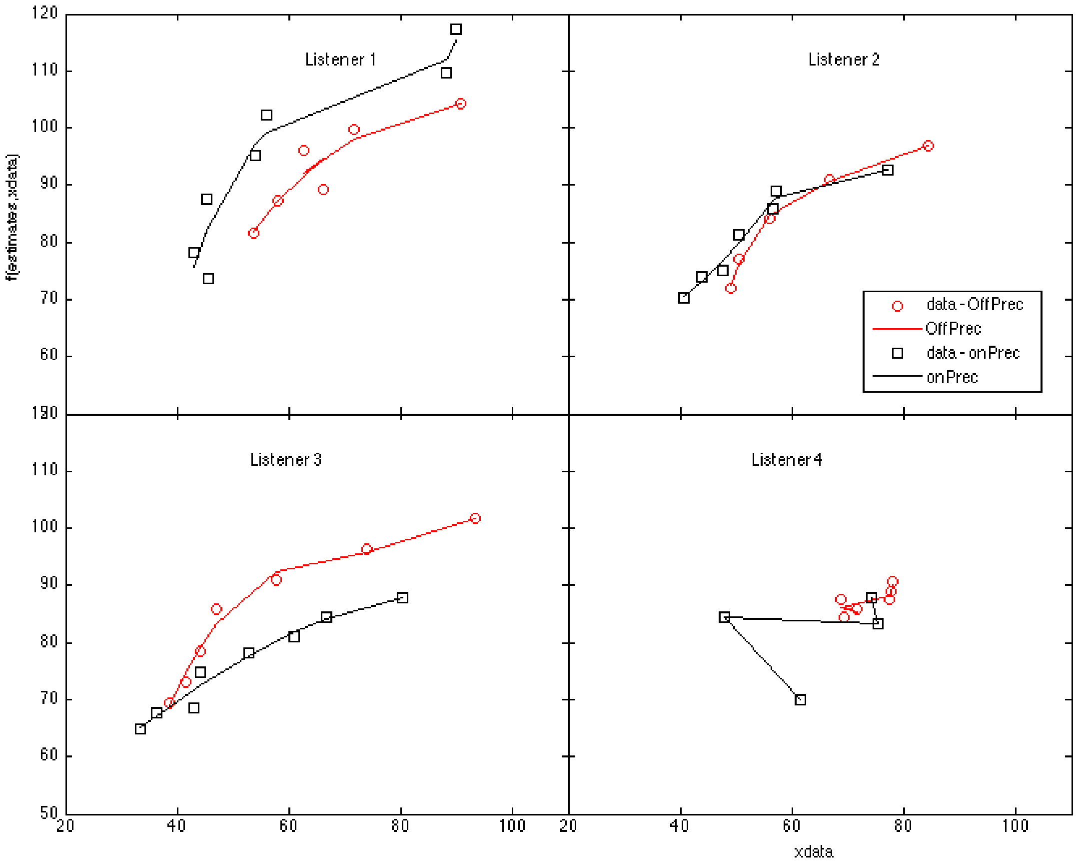


Figure 13: Results of XP1 plotted as BM I/O functions (symbols). Broken-stick model fits (lines) were applied to the functions in the on- (red) and off-frequency precursor (black) conditions.

The experimental data of L4 will be excluded from the following discussion as there is no clear BM I/O function for on- and off-frequency precursor visible.

According to our hypothesis based on MOCR gain reduction, activation of MOCR by the on-frequency precursor should induce a rightward shift of the I/O function for the off-frequency precursor. In contrast, the results in Fig. 13 show either a *left*ward shift (L1) or no shift at all (L2). One reason might be the constant target SPL for on- and off-frequency precursor. Only the results for L3 verify our hypothesis.

Specifically, the on-frequency precursor masks the shortest target by about 10 dB while the off-frequency precursor does not mask this target. Thus, having a constant target SPL implies that the sensation level[4] (SL) of the shortest target is higher in the off-

[4]The sensation level of a signal refers to its level in dB above its detection threshold.

than in the on-frequency condition. Consequently, higher masker levels are required to mask the target in off- than in on-frequency precursor conditions, as observed in Fig. 13. Having a constant SL rather than a constant SPL for the shortest target might solve the problem of the shifted off-frequency precursor curve. This will be done in the next experiment (XP2).

As mentioned above, the audibility of the cue signal at long target durations had to be clarified. To do so, a follow-up experiment was conducted at the ARI in December 2014. In this experiment, to ensure that the cue signal remains audible in all conditions, the cue signal duration was extended to the duration of precursor plus masker. In other words, the cue signal was gated on with the precursor onset and off with the masker offset. The goal was to provide a higher SL of the cue signal (detection thresholds decrease as signal duration increases, due to temporal integration in the auditory system) together with a better emphasis of the target onset. A constant SPL of 30 dB was used. In addition, the role of the cue signal on the variability was again assessed by measuring FDMCs with and without cue signal. Listeners were trained at least 2h in conditions with and without cue signal. Results for four listeners showed that first, thresholds with and without cue signal were very similar. Second, the variability was generally similar across conditions. In other words, the cue signal did not reduce the variability. It seems that training helped reducing the variability. Therefore, it was decided to measure FDMCs without cue signal in the main experiment described below.

2.7 XP2: Fixed-Duration Masking Curves without Cue Signal

The settings of XP2 are mentioned in Fig. 10 and Table 1. This experiment was conducted in February and March 2015.

2.7.1 Design

At first a pretest has been executed to determine the absolute threshold of the target and the masked threshold produced by the precursor. In this experiment the level of the shortest target was set to 10 dB SL for both the on- and off-frequency precursors. Precisely, the target SPL in the on-frequency condition was set to 10 dB above the masked threshold of the target in presence of the on-frequency precursor. The target SPL in the off-frequency condition was set to 10 dB above the threshold of the target in quiet.

Ten listeners participated. They were trained for five to seven hours, depending on the performance of the listener, on a subset of conditions. The main experiment took about eight hours to complete.

A problem occurred for the shortest target durations, despite training. Some listeners could not detect the target in presence of the off-frequency masker and thresholds remained unstable. For those listeners, an additional training period was provided (training of the item about 8-10 times). This additional training did not solve the problem for all listeners. For listeners with no improvement, some new long target durations of 13 ms, 15 ms, and 17 ms, were added to get enough measurement points for the off-frequency masker curve.

2.7.2 Results and Discussion

The results for on- (left column) and off-frequency precursor (right column) for one typical listener are shown as FDMCs in Fig. 14.

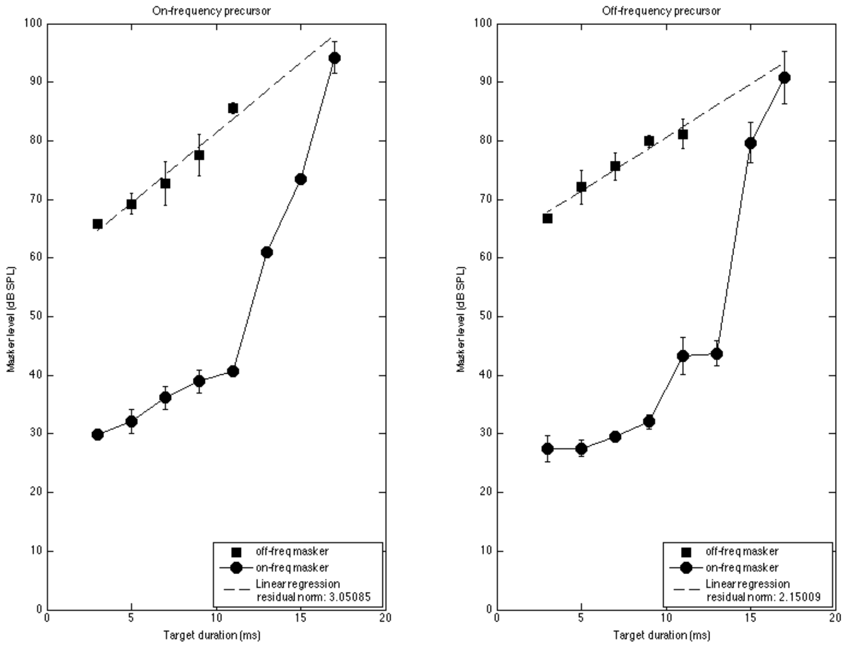


Figure 14: Result of XP2 for one listener plotted as FDMCs.

Again a linear regression (dashed line) was used for the off-frequency masker conditions to derive the BM I/O function for all target durations. For this listener the results of the measured thresholds for the off-frequency masker condition are linear, as expected.

The standard deviation for every measured threshold was limited to six for all results. For this listener there are still some problems at long target durations. Nevertheless, the limit was not exceeded for all conditions and the results could be taken into account.

The BM I/O functions derived for all listeners are shown in Fig. 15. Again all functions were fitted with the broken-stick model.

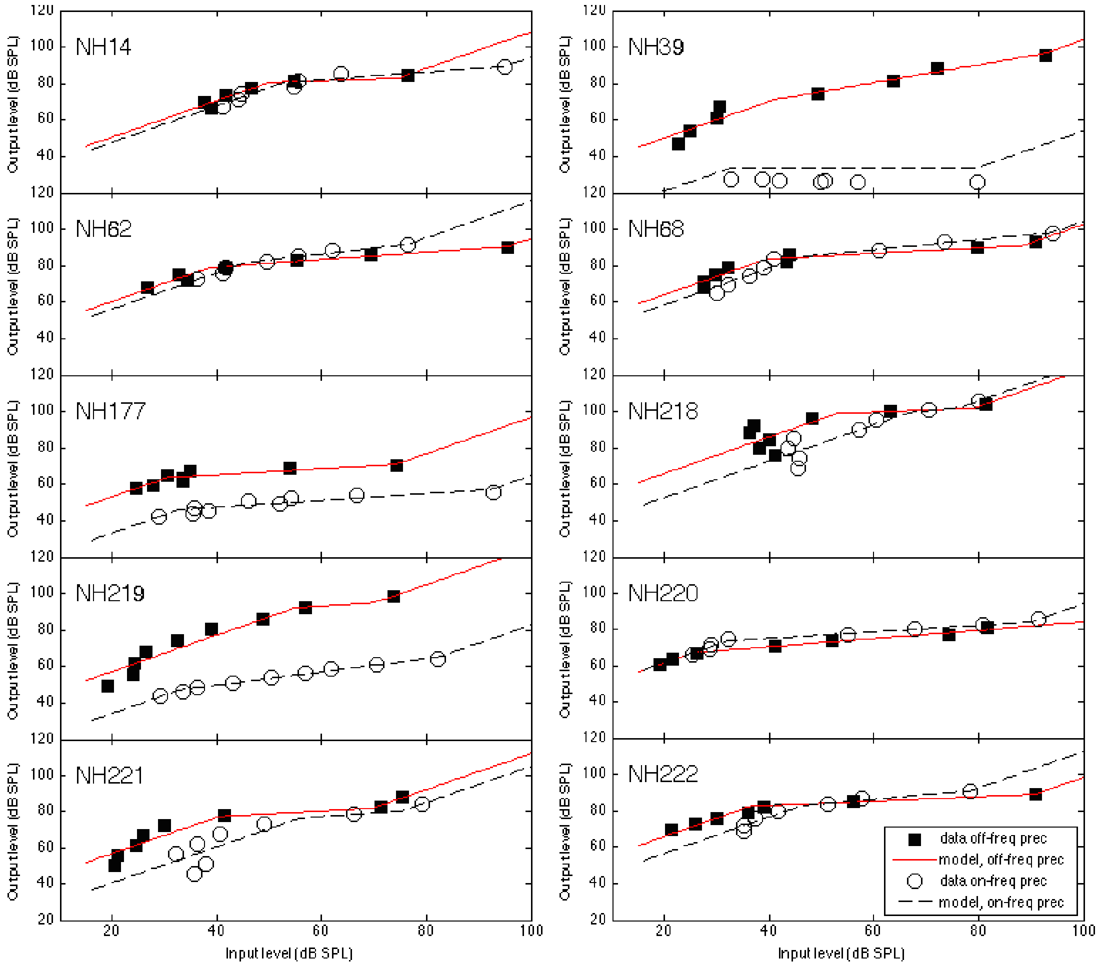


Figure 15: Results of XP2 plotted as I/O functions (symbols) for all listeners. Broken-stick model fits (lines) were applied to the functions in the on- (circles) and off-frequency precursor (squares) conditions.

The red line represents the measured BM I/O function without activation of MOCR whereas the black dashed line represents the measured BM I/O function with activated MOCR. As explained in Sec. 1.2 the aim of this study was to verify the changes in BM I/O function when the MOCR is activated. According to the hypothesis a decrease in G and an increase in BP_1 are expected. As MOCR only effects low level sounds BP_2 should not be affected. Therefore, there must be an increase in c. In other words, the compression of the BM is expected to decrease.

The broken-stick model parameters for all listeners are listed in Table 2.

	off-frequency Precursor				**on-frequency Precursor**			
	G	c	BP1	BP2	G	c	BP1	BP2
NH14	31	0.09	50	75	**28**	**0.21**	**53**	95
NH39	30	0.49	41	93	**1**	0	33	80
NH62	40	0.2	39	96	**36**	**0.33**	**45**	77
NH68	44	0.15	40	88	**39**	**0.28**	**46**	**94**
NH177	63	0.17	31	74	**13**	**0.19**	**33**	93
NH218	46	0.11	53	79	**33**	**0.28**	**67**	**77**
NH219	37	0.19	55	70	**14**	**0.36**	33	82
NH220	42	0.23	26	101	42	0.18	**32**	90
NH221	37	0.16	41	70	**21**	**0.22**	**56**	**76**
NH222	46	0.12	37	91	**37**	**0.25**	**46**	78

Table 2: Broken-stick model parameters estimated for all listeners. Bold values indicate changes that are compatible with the MOCR hypothesis.

A reduction in G occurred nearly for all listeners except for NH220. For half of the listeners this reduction is very strong whereas for all others it is rather weak. All listeners except NH39 and NH219 showed an increase in BP_1. Although no change in BP_2 was expected, BP_2 increased or decreased for seven listeners. The reasons for theses changes are unclear but it could be due to the fact that only few points have been obtained for $L_{in} > 80$ dB. Since BP_2 is supposed to be located around 80 dB (Jennings et al., 2009; Krull and E. A. Strickland, 2008), the present estimates of BP_2 might not be very representative. The most important observation is the increase in c and therefore a decrease in compression. This is observed for all listeners except for NH39 and NH220. As these listeners also reveal problems with BP_1 the hypothesis for MOCR could not be proved for these two listeners.

Overall, the results in Fig. 15 are compatible with those in Fig. 7. That is, there is a trend for an increase in c (i.e. a reduction in basilar membrane compression) when efferent effects like the MOCR are activated. The effect is different for every listener, though. For NH177, NH218, NH219 and NH221 the effect is very strong, whereas for NH14, NH62, NH68, NH222 it is very weak.

3 Conclusion

The aim of this study was to investigate the reduction in BM compression in humans when efferent effects like the MOCR are activated. In humans an indirect method must be used to measure BM I/O function. For this study the FDMC method was chosen as it results in reliable results and, if parameters are well chosen, it is not affected by efferents.

To estimate BM I/O functions with and without MOCR a precursor signal was presented before the masker and target stimuli. Conditions with an off-frequency precursor (i.e. remote from the target frequency) provide an efferent-free estimation of the I/O function. Conditions with an on-frequency precursor (i.e. centered at the target frequency) provide an estimation of the I/O function with efferents.

A reduction in compression of the BM has been hypothesized in this study. The results in Fig. 15 tend to confirm this hypothesis, although variability across listeners is observed.

References

Backus, B. C. and Jr. J. J. Guinan (2006). "Time course of the human medial olivocochlear reflex". In: *Journal of the Acoustical Society of America* 119 (cit. on p. 10).

Cooper, N. P. and J. J. Guinan (2006). "Efferent-mediated control of basilar membrane motion". In: *The Journal of Physiology* 576.1, pp. 49–54 (cit. on p. 11).

Glasberg, B. R. and B. C. J. Moore (1990). "Derivation of auditory filter shapes from notched-noise data". In: *Hearing Research* 47, pp. 103–138 (cit. on p. 16).

Guinan Jr., John J. (2006). "Olivocochlear efferents: Anatomy, physiology, function, and the measurement of efferent effects in humans". In: *Ear and Hearing*, pp. 589–607 (cit. on p. 14).

James, A. L., R. J. Mount, and R. V. Harrison (2002). "Contralateral supression of DPOAE measured in real time". In: *Clin. Otolaryngol* 27 (cit. on p. 10).

Jennings, Skyler G., Elizabeth A. Strickland, and Michael G. Heinz (2009). "Precursor effects on behavioral estimates of frequency selectivity and gain in forward masking". In: *The Journal of the Acoustical Society of America* 125.4, pp. 2172–2181 (cit. on pp. 11, 26).

Krull, V. and E. A. Strickland (2008). "The effect of a precursor on growth of forward masking". In: *Journal of the Acoustical Society of America* 123 (cit. on pp. 11, 12, 26).

Levitt, H (1971). "Transformed up-down methods in psychoacoustics". In: *The Journal of the Acoustical Society of America* 49.2, pp. 467–477 (cit. on p. 17).

Lilaonitkul, Watjana and John J. Guinan Jr. (2009). "Reflex control of the human inner ear: A half-octave offset in medial efferent feedback that is consistent with an efferent role in the control of masking". In: *Journal of Neurophysiology* 101 (cit. on p. 14).

Moore, Brian C. J. (2012). *An introduction to the psychology of hearing*. Sixth. Emerald (cit. on p. 11).

Neff, Donna L. (1985). "Stimulus parameters governing confusion effects in forward masking". In: *The Journal of the Acoustical Society of America* 78.6, pp. 1966–1976 (cit. on pp. 15, 16).

Nelson, D. A., A. C. Schroder, and M. Wojtczak (2001). "A new procedure for measuring peripheral compression in normal-hearing and hearing-impaired listeners". In: *Journal of the Acoustical Society of America* 110 (cit. on p. 13).

Oxenham, Andrew J. and Sid P. Bacon (2003). "Cochlear Compression: Perceptual Measures and Implications for Normal and Impaired Hearing". In: *Lippincott Williams & Wilkins* (cit. on pp. 6–8, 11).

Oxenham, Andrew J. and Christopher J. Plack (1997). "A behavioral measure of basilar-mebrane nonlinearity in listeners with normal and impaired hearing". In: *Journal of the Acoustical Society of America* 101 (cit. on pp. 12, 13).

Ren, Tianying (2002). *Longitudinal pattern of basilar membrane vibration in the sensitive cochlea*. Online: `http://www.pnas.org/content/99/26/17101/F2.expansion.html` (zuletzt besucht am 15. Juli 2015) (cit. on p. 7).

Roverud, E. and E. A. Strickland (2010). "The time course of cochlear gain reduction measured using a more efficient psychophysical technique". In: *Journal of the Acoustical Society of America* 128 (cit. on pp. 10–12).

Strickland, Elizabeth A. (2008). "The relationship between precursor level and the temporal effect". In: *Journal of the Acoustical Society of America* 123 (cit. on pp. 9–12).

Yasin, Ifat, Vit Drga, and Christopher J. Planck (2013). "Estimating peripheral gain and compression using fixed-duration masking curves". In: *Journal of the Acoustical Society of America* 133.6 (cit. on pp. 11, 13).

Yasin, Ifat and Christopher J. Plack (2003). "The effects of a high-frequency suppressor on tuning curves and derived basilar-membrane response functions". In: *The Journal of the Acoustical Society of America* 114.1, pp. 322–332 (cit. on p. 9).