Lisa Kristina Kerle

Influence of the form factor on the pitch of glides

Bachelorarbeit

eingereicht an der

Universität für Musik und darstellende Kunst Graz

Im Rahmen der Lehrveranstaltung:

17.0013 Musikalische Akustik 01 SE

Betreuer

o.Univ.Prof. Mag.art. DI Dr. Robert Höldrich

Institut für Elektronische Musik und Akustik – IEM

Graz, September 2018

Abstract

This thesis deals with the investigation of the so-called *glides*. Glides are short sound impulses that change their instantaneous frequency either in ascending or descending direction and are characterized by their duration, by the frequency span the glide runs through and by the arithmetic mean of the frequency span.

But this frequency span is not perceived for each glide. For glides with sufficiently small durations and frequency spans only one single pitch is evoked by the glide. This perceived pitch is not automatically the frequency at the temporal center of the glide. The larger the duration-bandwidth-product of the glide is, the higher is the tendency to perceive a pitch near the final frequency of the glide.

To investigate the perceived pitch a listening experiment is carried out in which various glides with different durations, bandwidths and arithmetic center frequencies are tested.

But before the listening experiment is described, a short review of the existing literature and the found facts will be given.

Subsequently the listening experiment is described in which various form factors were examined. These form factors influence the course of the instantaneous frequency.

For various form factors two glides with identical duration, arithmetic center frequency and bandwidth but different direction are compared to each other. It is the main objective of this thesis to find a certain form for the up-glide and the down-glide to perceive the same pitch for both glides.

Afterwards a statistical analysis is carried out to show the significance of the achieved results.

In conclusion further investigations are explained that could be carried out based on the obtained outcome.

Deutsche Zusammenfassung

Wie der Titel schon andeutet, beschäftigt sich die folgende Arbeit mit den sogenannten Glides. Glides sind kurze Tonimpulse, welche ihre Momentanfrequenz in aufsteigender oder absteigender Richtung verändern und durch ihre Dauer, die Bandbreite des durchlaufenen Frequenzbereichs und den arithmetischen Mittelwert des Frequenzbereichs definiert sind.

Bereits vorangegangene Untersuchungen zeigen, dass bei hinreichend kleinen Dauern und Frequenzbereichen keine Frequenzänderung wahrgenommen wird, sondern lediglich eine einzelne Tonhöhe.

Um diese wahrgenommene Tonhöhe genauer zu erforschen, wird ein Experiment mit jenen Glides durchgeführt.

Bevor das Experiment jedoch genauer erläutert wird, gibt diese Arbeit noch einen kurzen Einblick in die vorhandene Literatur und die bereits durchgeführten Experimente.

Anschließend folgt eine Beschreibung des Experiments, bei welchem je zwei Glides miteinander verglichen werden, welche in der Dauer, dem Umfang der Frequenzänderung und der arithmetischen Mittenfrequenz ident sind, sich allerdings in der Richtung der Frequenzänderung voneinander unterscheiden. Es ist nun das Ziel, die Form der Glides, also den Verlauf der Momentanfrequenz, so zu verändern, dass für den Up-Glide und den Down-Glide die selbe Tonhöhe wahrgenommen wird. Es werden unterschiedliche Bedingungen untersucht und für jede Bedingung ein Formfaktor ermittelt, bei welchem die Tonhöhe von Up-Glide und Down-Glide übereinstimmt.

Es folgt eine statistische Analyse der Ergebnisse, um die Signifikanz der erhaltenen Werte zu überprüfen.

Abschließend wird der Leser auf weiterführende Experimente hingewiesen.

Table of Contents

List of Figures

List of Tables

1 Introduction

1.1 What is a Glide?

Fundamentally this thesis deals with the investigation of short sound impulses that change their instantaneous frequency either in ascending or descending direction, the so-called *glides*. Glides with ascending frequency are called *up-glides* and glides with descending frequency are called *down-glides*.

Those glides are mainly characterized by duration *T*, arithmetic center frequency f_{ma} and bandwidth *B*. Duration means the period of time between the onset and the offset of the glide. The bandwidth is defined as the difference between the highest and the lowest frequency the glide runs through, in other words the transition span, and the arithmetic mean of those two frequencies describes the arithmetic center frequency f_{ma} .

1.2 Pitch of Glides

For sufficient short durations and small bandwidths humans are not able to recognize a frequency change while listening to the glide but perceive a clear pitch (cf. Nabelek et al., 1970: p.539, The Journal of the Acoustical Society of America, Vol. 48). As explored in previous studies this perceived pitch cannot readily be determined without any investigations.

Considering for example linear glides where the instantaneous frequency changes linearly with time, the up-glide rather evokes a higher pitch than the down-glide although not only bandwidth and arithmetic center frequency but also duration is identical (cf. Nabelek et al., 1970: p.539, The Journal of the Acoustical Society of America, Vol. 48). As both the up-glide and the down-glide have according to identical duration and bandwidth the same glide rate and therefore the same center frequency, which means the instantaneous frequency at the temporal center of the glide, it seems obvious that not the frequency at the temporal center of the glide is perceived but a frequency located somewhere in the rear area of the glide.

This leads to the assumption that glides with same settings in bandwidth, arithmetic center frequency and duration but different direction do not cause the same pitch automatically. Considering glides with a sufficiently large duration-bandwidth-product, the resulting pitch tends to be somewhere in the rear half of the glide (cf. Nabelek et al., 1970: p.539, The Journal of the Acoustical Society of America, Vol. 48). Therefore it would be interesting to create not only linear glides but also glides with a nonlinear frequency change and to examine their impact on pitch perception.

1.3 Mathematical Structure of Glides

The description now goes into more detail to explain the mathematical structure of linear and nonlinear glides using formulas and graphical representations.

1.3.1 Linear Glides

First linear glides are investigated which are described by duration *T*, center frequency f_0 and bandwidth *B*.

Generally it is assumed that the glide runs from $\frac{-T}{2}$ to $\frac{T}{2}$. Bandwidth *B* is the difference between the lowest frequency f_{low} and the highest frequency f_{high} the glide passes through. As only glides with rising or falling instantaneous frequency are examined, the lowest and highest frequencies correspond to the start and end frequencies of the glide.

Considering linear glides, the center frequency f_0 is the instantaneous frequency at time *t=0* which is in other words described by the arithmetic mean of f_{high} and f_{low} . So for linear glides the arithmetic center frequency f_{ma} is identical to the center frequency f_0 .

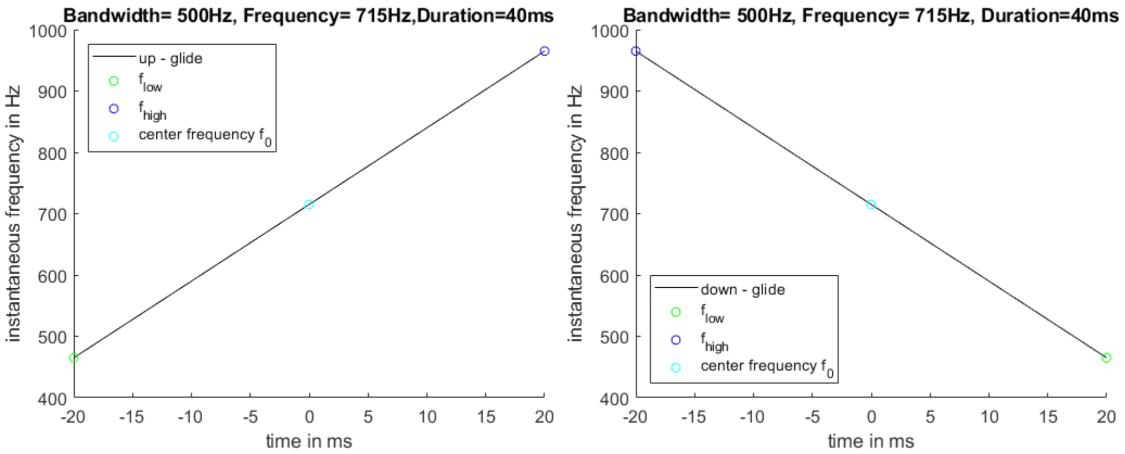

figure 1: Instantaneous frequency of linear up-glide (left) and down-glide (right) with same bandwidth, center frequency and duration

According to the linear equation $y=kx+d$ the instantaneous frequency $f_i(t)$ of linear glides with glide rate a in [Hz/s] is given as follows:

- up-glide: $f_{i,up}(t)=f_0+at$
- down-glide $f_{i,down}(t)=f_0-at$

As mentioned above the value of the instantaneous frequency at time *t=0* is the center frequency f_0 which is consistent with the structure of the equations above.

To determine the slope of linear glides, which is called the glide rate a , duration *T* and bandwidth *B* are used.

Considering the figure above (*figure 1*) it becomes clear that the instantaneous frequency $f_{i,up}(t)$ of the up-glide deviates at time $t=\frac{-T}{2}$ from the value of the center frequency f_0 by $\frac{-B}{2}$ and the instantaneous frequency $f_{i,down}(t)$ of the down-glide deviates at the same time from f_0 by $\frac{+B}{2}$, which leads to the following equations:

- up-glide: $f_{i,up}(\frac{-T}{2})=f_0+a\frac{-T}{2}=f_0-\frac{B}{2}$ → $a\frac{T}{2}=\frac{B}{2}$

- down-glide: $f_{i,down}\left(\frac{-T}{2}\right)=f_0-a\frac{-T}{2}=f_0+\frac{B}{2}$ → $a\frac{T}{2}=\frac{B}{2}$

Solving those equations yield the unknown glide rate a . This glide rate is therefore given by $a=\frac{B}{T}$ and makes it possible to formulate the instantaneous frequency of the up- and down-glide:

- up-glide: $f_{i,up}(t)=f_0+\frac{B}{T}t$
- down-glide: $f_{i,down}(t)=f_0-\frac{B}{T}t$

To describe linear glides, these instantaneous frequencies have to be integrated from $\frac{-T}{2}$ to $\frac{T}{2}$ and multiplied with 2π to get the instantaneous phase of the up-glide and the down-glide in [rad].

- up-glide: $\varphi_{i,up}(t)=2\pi\int f_{i,up}(t)dt=2\pi\left(f_0t+\frac{B}{T}\frac{t^2}{2}\right)$
- down-glide: $\varphi_{i,down}(t)=2\pi\int f_{i,down}(t)dt=2\pi\left(f_0t-\frac{B}{T}\frac{t^2}{2}\right)$

Finally linear glides can be fully described for $\frac{-T}{2}\leq t\leq\frac{T}{2}$ by the following equations:

- up-glide: $x_{up}(t)=\cos(\varphi_{i,up}(t))$
- down-glide: $x_{down}(t)=\cos(\varphi_{i,down}(t))$

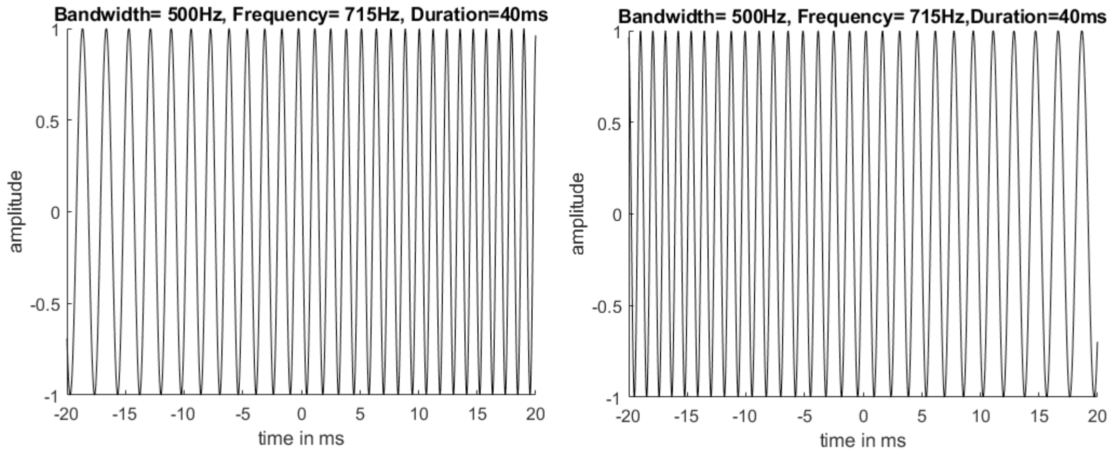

figure 2: Linear up-glide (left) and down-glide (right) with same bandwidth, center frequency and duration

1.3.2 Exponential Glides

The following paragraph deals with the investigation of nonlinear glides, especially exponential glides, whose instantaneous frequency changes exponentially with time.

For linear glides the instantaneous frequency at the temporal center *t=0* is the center frequency f_0 . Therefore it would be advantageous that this condition holds for exponential glides too, yielding

$$f_{i,\exp}(t)=f_0+A\left(e^t-1\right)$$

As in the linear case, exponential glides are also characterized by duration *T* as the glide runs from $\frac{-T}{2}$ to $\frac{T}{2}$ and the bandwidth *B*, described by the difference between the highest frequency f_{high} and the lowest frequency f_{low} of the glide which are again consistent to start and end frequency. But it is important to mention that for nonlinear glides the center frequency f_0 is not identical to the arithmetic center frequency f_{ma} anymore. As for nonlinear glides frequency does not change linearly with time, a frequency f_0 different from the arithmetic center frequency f_{ma} is located at time *t=0*. This fact is depicted in the graphic below.

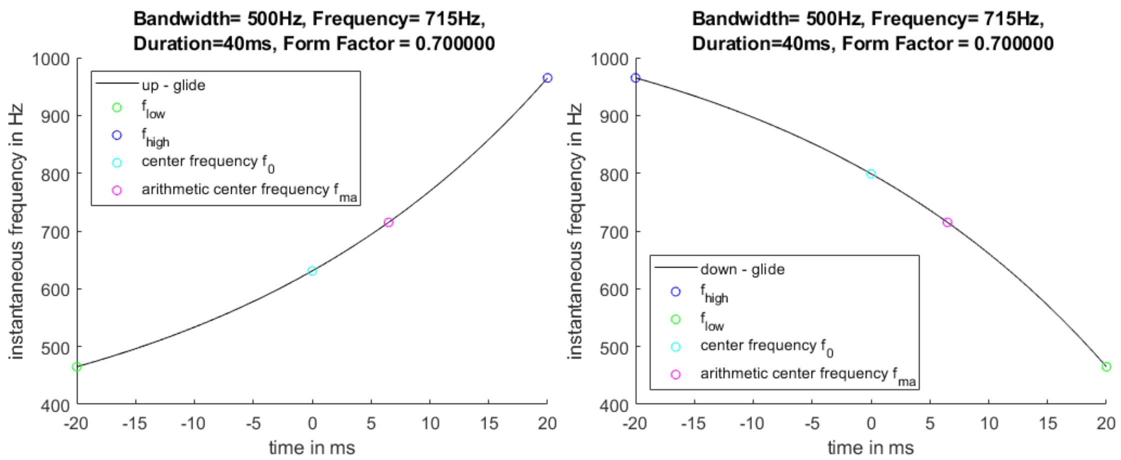

figure 3: Instantaneous frequency of exponential up-glide (left) and down-glide (right) with same bandwidth, arithmetic center frequency, duration and form factor

Apart from that it must be considered that up-glides and down-glides with same settings in duration, bandwidth and arithmetic center frequency differ in their center frequencies (instantaneous frequency at time *t=0*) as these center frequencies $f_{0,up}$ and $f_{0,down}$ are not the arithmetic mean of start and end frequency anymore as in the linear case.

In addition to the parameters mentioned above, another essential parameter plays an important role concerning the shape of nonlinear glides. This parameter is called the form factor α which is mentioned in the plot above and influences the curvature.

Combining the two different center frequencies $f_{0,up}$ and $f_{0,down}$ and this form factor leads to the following equations for the instantaneous frequency of exponential glides:

- up-glide: $f_{i,\exp,up}(t) = f_{0,up} + A\left(e^{\alpha t} - 1\right)$
- down-glide: $f_{i,\exp,down}(t) = f_{0,down} - A\left(e^{\alpha t} - 1\right)$

In order to create glides, whose shapes are independent of duration *T*, this form factor has to be normalized:

$$\alpha \rightarrow \frac{\alpha}{T/2}$$

To solve the equations above and to get the unknown parameters *A* and $f_{0,up}$ respectively $f_{0,down}$, the same approach as above is used.

The glide should run through the bandwidth *B* during the duration *T*, at which the instantaneous frequency $f_{i,\exp,up}$ at time $t=\frac{-T}{2}$ is f_{low} and the instantaneous frequency at time $t=\frac{T}{2}$ is f_{high} if the up-glide is considered which starts at the lowest frequency and ends at the highest frequency. Considering the down-glide, the instantaneous frequency $f_{i,\exp,down}$ at the starting point $t=\frac{-T}{2}$ is the highest frequency f_{high} and at the end point $t=\frac{T}{2}$ the value of $f_{i,\exp,down}$ is the lowest frequency f_{low} .

The values of f_{high} and f_{low} result from the following equations:

- $f_{high}=f_{ma}+\frac{B}{2}$
- $f_{low}=f_{ma}-\frac{B}{2}$

These conditions yield the following equations:

- up-glide:

(1) $f_{i,\exp,up}\left(\frac{-T}{2}\right)=f_{0,up}+A\left(e^{\frac{-\alpha}{T/2}\frac{T}{2}}-1\right)=f_{low}$

(2) $f_{i,\exp,up}\left(\frac{T}{2}\right)=f_{0,up}+A\left(e^{\frac{\alpha}{T/2}\frac{T}{2}}-1\right)=f_{high}$

- down-glide:

(3) $f_{i,\exp,down}\left(\frac{-T}{2}\right)=f_{0,down}-A\left(e^{\frac{-\alpha}{T/2}\frac{T}{2}}-1\right)=f_{high}$

(4) $f_{i,\exp,down}\left(\frac{T}{2}\right)=f_{0,down}-A\left(e^{\frac{\alpha}{T/2}\frac{T}{2}}-1\right)=f_{low}$

Multiplying the first equation (1) with the factor (-1) and adding this to equation (2) leads to the demanded parameter *A* for up-glides:

$$-f_{i,\exp,up}\left(\frac{-T}{2}\right)+f_{i,\exp,up}\left(\frac{T}{2}\right)=-A\left(e^{\frac{-\alpha}{T/2}\frac{T}{2}}-1\right)+A\left(e^{\frac{\alpha}{T/2}\frac{T}{2}}-1\right)=-f_{low}+f_{high}=B$$

$$\rightarrow \quad A\left(e^{\frac{\alpha}{T/2}\frac{T}{2}}-e^{\frac{-\alpha}{T/2}\frac{T}{2}}\right)=B \quad \rightarrow \quad A=\frac{B}{2\sinh\left(\frac{\alpha}{T/2}\frac{T}{2}\right)}=\frac{B}{2\sinh(\alpha)}$$

For the down-glide the fourth equation (4) is multiplied with the factor (-1) and is then added to equation (3) which leads to the same solution for *A*.

The instantaneous frequency of nonlinear glides is therefore expressed by the following equations:

- up-glide: $f_{i,\exp,up}(t)=f_{0,up}+\frac{B}{2\sinh(\alpha)}(e^{\frac{\alpha}{T/2}t}-1)$
- down-glide: $f_{i,\exp,down}(t)=f_{0,down}-\frac{B}{2\sinh(\alpha)}(e^{\frac{\alpha}{T/2}t}-1)$

By using this form factor α as mentioned above, the shape of the exponential glides is influenced. For $\alpha\approx 0$ a linear glide is created, increasing α results in a more curved shape and for negative α values the curvature is changing its direction. What is important to mention is that the form factor must not be equal to zero as $\sinh(0)=0$, which would lead to a division by zero.
For fixed parameters f_{high} and f_{low} , the unknown frequencies at *t=0* $f_{0,up}$ respectively $f_{0,down}$ can be calculated with assistance of the following equations:

- up-glide: $f_{ma}-\frac{B}{2}=f_{low}=f_{i,\exp,up}\left(\frac{-T}{2}\right)=f_{0,up}+\frac{B}{2\sinh(\alpha)}\left(e^{\frac{-\alpha}{T/2}\frac{T}{2}}-1\right)$

$$\rightarrow \quad f_{0,up}=f_{ma}-\frac{B}{2}-\frac{B}{2\sinh(\alpha)}(e^{-\alpha}-1)$$

- down-glide: $f_{ma}+\frac{B}{2}=f_{high}=f_{i,\exp,down}\left(\frac{-T}{2}\right)=f_{0,down}-\frac{B}{2\sinh(\alpha)}\left(e^{\frac{-\alpha}{T/2}\frac{T}{2}}-1\right)$

$$\rightarrow \quad f_{0,down}=f_{ma}+\frac{B}{2}+\frac{B}{2\sinh(\alpha)}(e^{-\alpha}-1)$$

Finally the instantaneous frequency of exponential glides can be fully described by the parameters *B*, *T*, f_{ma} and α :

- up-glide: $f_{i,\exp,up}(t)=f_{ma}-\frac{B}{2}-\frac{B}{2\sinh(\alpha)}(e^{-\alpha}-1)+\frac{B}{2\sinh(\alpha)}\left(e^{\frac{\alpha}{T/2}t}-1\right)$

  → $f_{i,\exp,up}(t)=f_{ma}-\frac{B}{2}\left(1+\frac{e^{-\alpha}-1}{\sinh(\alpha)}-\frac{e^{\frac{\alpha}{T/2}t}-1}{\sinh(\alpha)}\right)=f_{ma}-\frac{B}{2}\left(1+\frac{e^{-\alpha}-e^{\frac{\alpha}{T/2}t}}{\sinh(\alpha)}\right)$

- down-glide: $f_{i,\exp,down}(t)=f_{ma}+\frac{B}{2}+\frac{B}{2\sinh(\alpha)}(e^{-\alpha}-1)-\frac{B}{2\sinh(\alpha)}\left(e^{\frac{\alpha}{T/2}t}-1\right)$

  → $f_{i,\exp,down}(t)=f_{ma}+\frac{B}{2}\left(1+\frac{e^{-\alpha}-1}{\sinh(\alpha)}-\frac{e^{\frac{\alpha}{T/2}t}-1}{\sinh(\alpha)}\right)=f_{ma}+\frac{B}{2}\left(1+\frac{e^{-\alpha}-e^{\frac{\alpha}{T/2}t}}{\sinh(\alpha)}\right)$

To get the instantaneous phase of the exponential glides, the instantaneous frequency is integrated from $\frac{-T}{2}$ to $\frac{T}{2}$ and multiplied with 2π to get the instantaneous phase in [rad].

- up-glide: $\varphi_{i,\exp,up}(t)=2\pi\int f_{i,\exp,up}(t)\,dt=2\pi\left(f_{0,up}t+A\left(\frac{e^{\frac{\alpha}{T/2}t}}{\frac{\alpha}{T/2}}-t\right)\right)$

- down-glide: $\varphi_{i,\exp,down}(t)=2\pi\int f_{i,\exp,down}(t)\,dt=2\pi\left(f_{0,down}t-A\left(\frac{e^{\frac{\alpha}{T/2}t}}{\frac{\alpha}{T/2}}-t\right)\right)$

This phase is now used to express the exponential glides for $\frac{-T}{2}\leq t\leq\frac{T}{2}$ as follows:

- up-glide: $x_{\exp,up}(t)=\cos(\varphi_{i,\exp,up}(t))$
- down-glide: $x_{\exp,down}(t)=\cos(\varphi_{i,\exp,down}(t))$

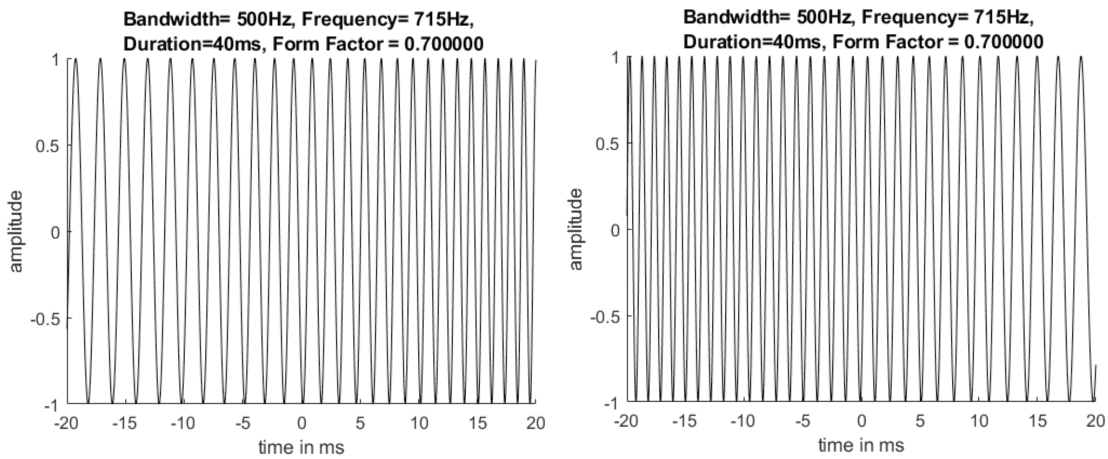

figure 4: Exponential up-glide (left) and down-glide (right) with same bandwidth, arithmetic center frequency, duration and form factor

If we have a look at these exponential glides, it becomes clear that the form factor α plays an important role in the investigation of the pitch. Now the question arises if it is possible to find an adequate form factor α at which the up-glide and the down-glide evoke the same pitch.

1.4 Objective and Content of this Thesis

It is the main objective of this thesis to investigate linear and nonlinear glides, especially exponential glides with different duration, bandwidth and arithmetic center frequency to find an appropriate form factor at which the up-glide and the down-glide are perceived to have the same pitch.
To achieve this aim a listening experiment is carried out in which different form factors are tested and to conclude this thesis, a statistical evaluation will be performed to show the significance of the achieved results.

The main part of this bachelor thesis will be the listening experiment to investigate the perceived frequency of up-glides and down-glides for various conditions that differ in bandwidth, duration and arithmetic center frequency.
In chapter 2 I would like to discuss some of the existing literature which deals with glides and especially with the evoked pitch of glides to give the reader a short introduction to the topic.
In chapter 3 the listening experiment will be described in detail. As this listening experiment works with the method of constant stimuli, a pretest has to be made to get suitable values that should be tested. So before going into detail about the listening experiment itself, the process of preparing the listening experiment will be explained. In this part not only the pretests will be mentioned but also a description of the used software will be given.
After presenting the analysis of this pretest, the procedure of the main experiment is described step by step. This means that some basic information about the experiment will be given and in between some graphical representations will be inserted to support the descriptions.
Unfortunately the first experiment did not yield a satisfying outcome. As a result of that a second listening experiment with modified settings had to be carried out in order to achieve a more useful performance. These modifications and the carrying out of the second experiment will be mentioned as well as the outcome of this second part.

Finally the results of the two listening experiments will be combined, the final outcome will be discussed and the statistical analysis presented.
Furthermore a little overview will be given for which investigations those results will be helpful and which studies could be carried out in addition to this listening experiment.

2 Perception of Glides – A Literature Review

Before starting with the main part of this thesis, the listening experiment, the reader should first get a short overview about what has already been researched on this topic. Therefore some essential papers will be mentioned in this chapter as the listening experiment is based on those previous investigations.

A big part of this already done research deals with detection and discrimination of linear and nonlinear glides. In addition to that the pitch of linear glides is investigated.

2.1 Linear Glides

2.1.1 Detection and Discrimination of Glides in Frequency

First it would be interesting to investigate subjects' aptitude in detecting and discriminating glides in frequency. An essential paper that deals with this topic is *Detection and discrimination of gliding tones as a function of frequency transition and center frequency* from Madden and Fire (1996).

To investigate subjects' performance in detecting linear glides, a two-alternative, forced-choice task was carried out in which glides are compared to stimuli that did not change in frequency. However for the discrimination task, glides were compared to stimuli that increase either by 250Hz or 500Hz. In both experiments different center frequencies and in addition to that as well roved and nonroved conditions were tested. Frequency roving means that the arithmetic center frequency of the glides is different for each trial.

It was now on the subjects to identify glides which increased in pitch (detection) or increased more in pitch (discrimination). The result of this two-alternative, forced-choice task yields the thresholds for which detection and discrimination were possible for different center frequencies. The results are presented below. Thresholds are given as a proportion of the equivalent rectangular bandwidth ERB. The equivalent rectangular bandwidth of a filter describes the bandwidth of a rectangular filter with same peak transmission and total power for a white noise input as that filter (cf. Madden and Fire, 1996: p.3760, The Journal of the Acoustical Society of America, Vol. 100, cited from Moore 1989, p.334).

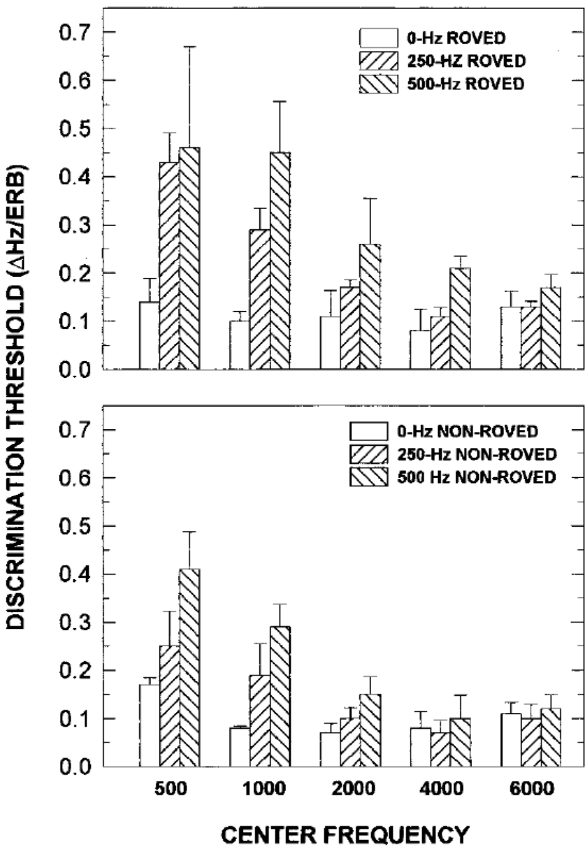

figure 5: Average discrimination thresholds for fixed standard transitions expressed as a proportion of the ERB and plotted as a function of center frequency. Error bars indicate one standard deviation. Madden/ Fire (1996)

In the picture it becomes obvious that thresholds in roved conditions are higher than nonroved thresholds (cf. Madden and Fire, 1996: p.3756, The Journal of the Acoustical Society of America, Vol. 100). Apart from that thresholds increase with standard transition for lower center frequencies but stay nearly constant for higher center frequencies (cf. Madden and Fire, 1996: p.3756, The Journal of the Acoustical Society of America, Vol. 100). Thresholds for glide discrimination depend on the center frequency for lower frequencies. This is shown in the figure above as thresholds increase as frequency decreases at center frequencies below 4kHz (cf. Madden and Fire, 1996: p.3756, The Journal of the Acoustical Society of America, Vol. 100). However glide detection thresholds vary only little with frequency (cf. Madden and Fire, 1996: p.3756, The Journal of the Acoustical Society of America, Vol. 100).

A further experiment was carried out in the paper *Detection and discrimination of frequency glides as a function of direction, duration, frequency span, and center frequency* by Madden and Fire (1997) to investigate the influence of direction, duration, frequency span and center frequency on the detection and discrimination of linear glides. In this experiment subjects were again asked to detect either the stimulus whose instantaneous frequency changed during the burst duration compared to a signal that was even in frequency or to identify the signal with higher glide rate compared to a signal that was swept across a fixed transition span.

Different durations, center frequencies, transition spans and directions were compared.

This experiment shows that center frequency only has little effect on the discrimination threshold (cf. Madden and Fire, 1997: p.2922, The Journal of the Acoustical Society of America, Vol. 102).

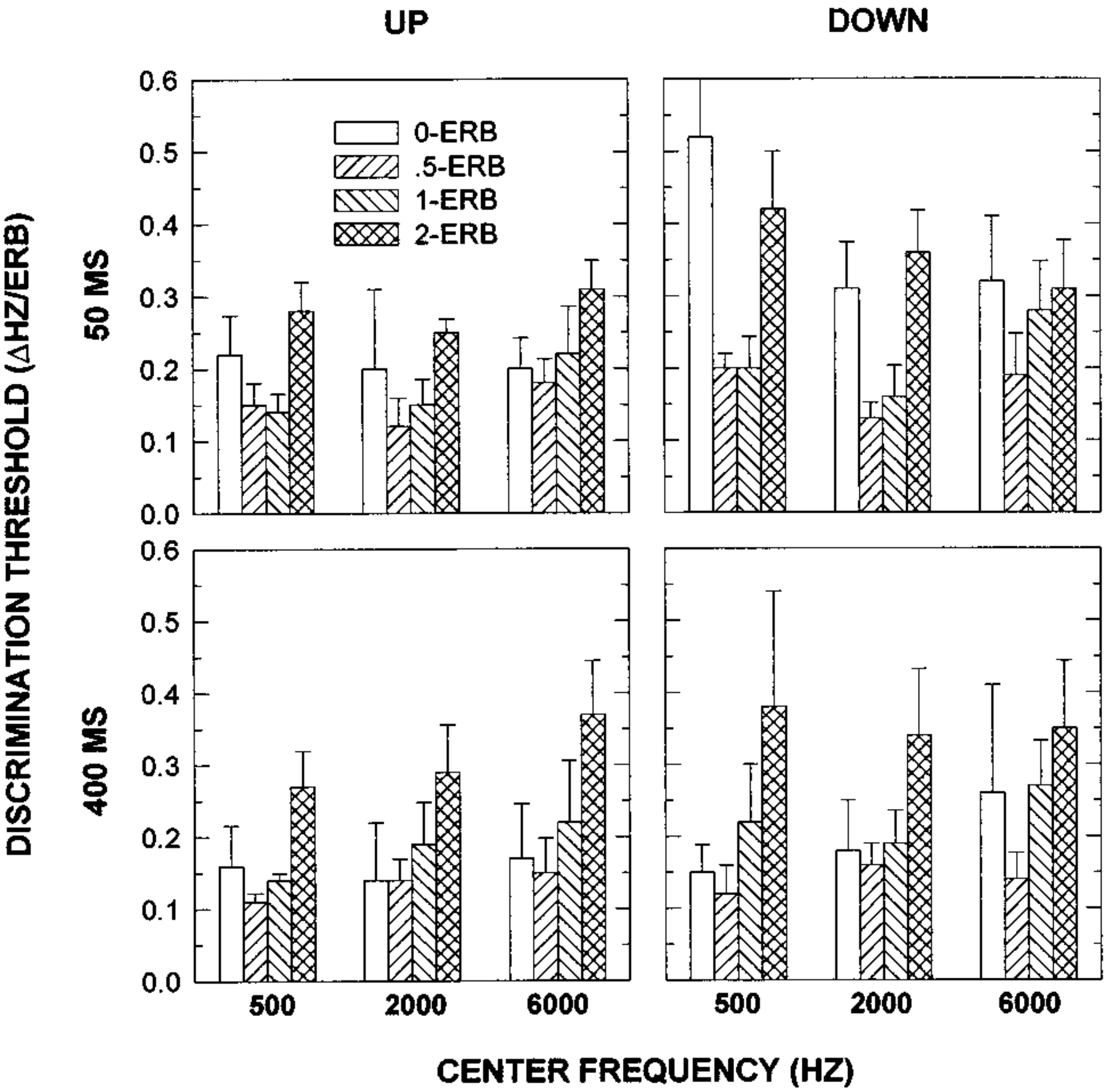

figure 6: Average discrimination thresholds for the various combinations of duration, direction, and frequency span, plotted as a function of center frequency. Thresholds are expressed as a proportion of the ERB at the signal center frequency. Error bars indicate one standard deviation from the mean. Madden/ Fire (1997)

This discrimination threshold is also nearly independent from glide duration and direction but for detection the direction was important for short durations as the 50-ms down-glides were more difficult to detect compared to the up-glides with same duration (cf. Madden and Fire, 1997: p.2923, The Journal of the Acoustical Society of America, Vol. 102).

The transition span did not influence thresholds for transition spans below 2-ERB but thresholds increased significantly for the 2-ERB span, except the 50-ms down-glides (cf. Madden and Fire, 1997: p.2923, The Journal of the Acoustical Society of America, Vol. 102).

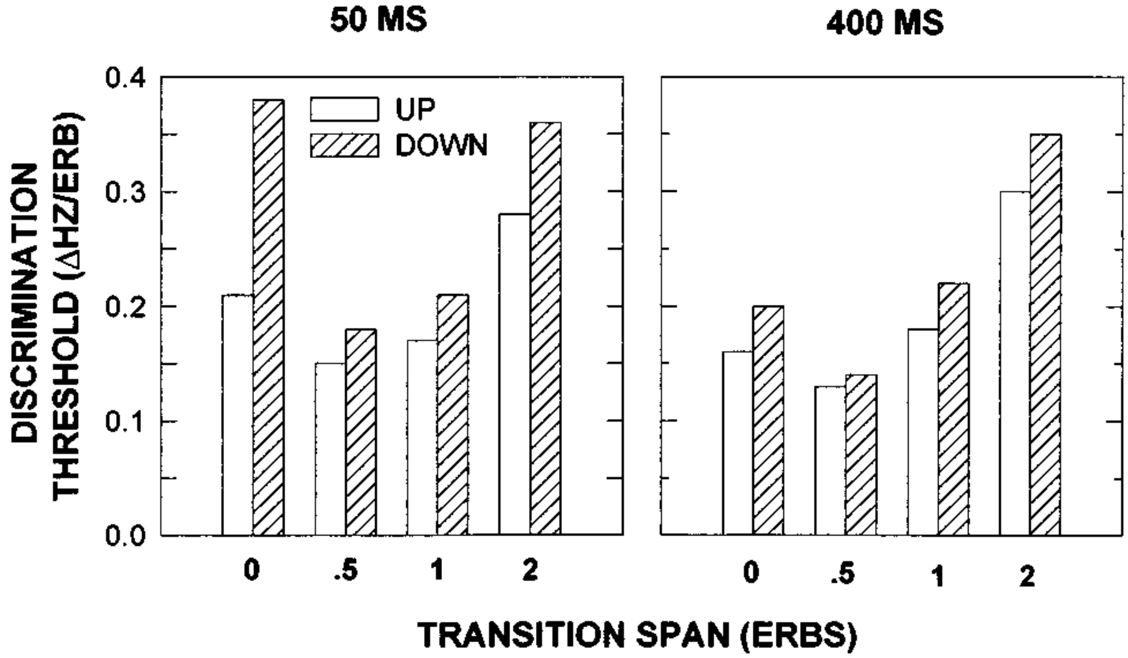

figure 7: Detection and discrimination thresholds for each signal duration, averaged across center frequency and plotted as a function of transition span. Madden/ Fire (1997)

2.1.2 Pitch of Linear Glides

This thesis is mainly about the comparison of the perceived pitch of up-glides and down-glides. Therefore the following paragraph will give a summary of what has already been found out concerning the pitch of linear glides.

An essential paper in this field is *Pitch of Tone Bursts of Changing Frequency* from Nabelek, Nabelek and Hirsh (1970).

One part of this paper deals with the presentation of glides whose transition span lasts during the whole burst duration. In case of small enough transition spans and durations a clear pitch was perceived by the subjects (cf. Nabelek et al., 1970: p.539, The Journal of the Acoustical Society of America, Vol. 48). This pitch is located near the middle of the burst but with a tendency to the final frequency of the glide (cf. Nabelek et al., 1970: p.539, The Journal of the Acoustical Society of America, Vol. 48). This effect seems to be stronger for rising than for falling glides and increases for higher duration-bandwidth-products (cf. Nabelek et al., 1970: p.539, The Journal of the Acoustical Society of America, Vol. 48).

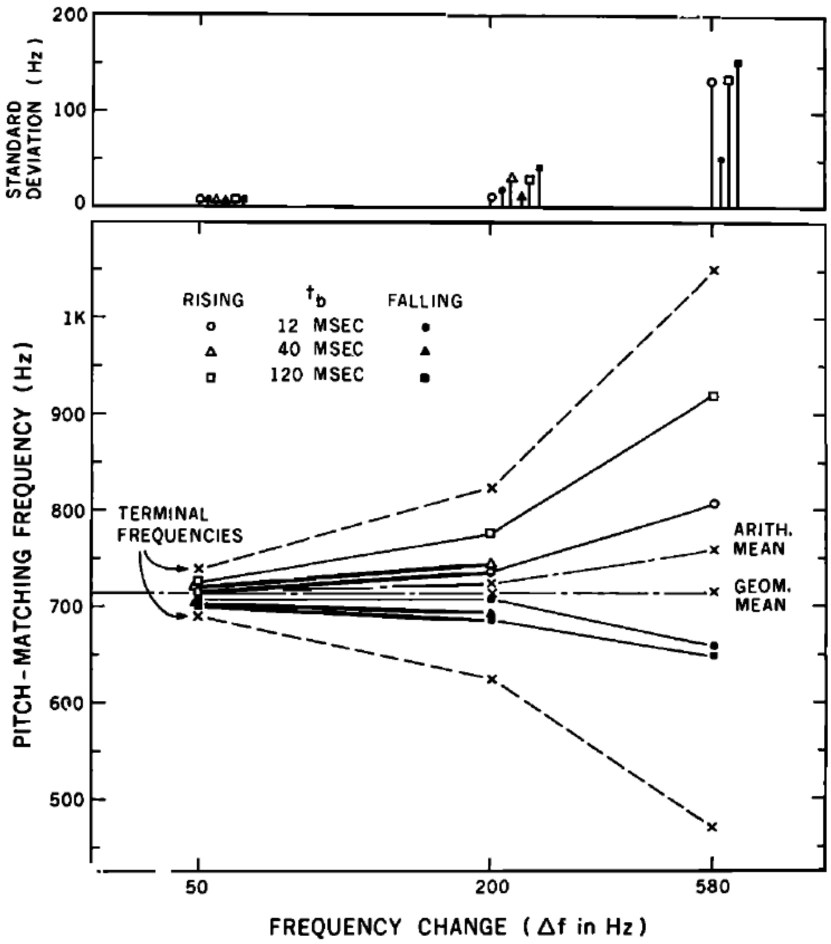

figure 8: Means and standard deviations of pitch-matching frequencies as a function of frequency change for bursts with 100% transition duration. Nabelek et al. (1970)

Apart from that an additional experiment was carried out by Nabelek et al. to find the maximum frequency extents at which fusion occurs and the minimum frequency change at which separation happens.

Different burst durations, transition durations, transition spans and rise and decay times of bursts were tested.
It is assumed that frequency changes that yield fusion or separation are similar for rising and falling glides and therefore only rising-frequency transitions were investigated (cf. Nabelek et al., 1970: p.544, The Journal of the Acoustical Society of America, Vol. 48). In this experiment transition spans were increased or decreased in small steps.
The longer the bursts gets the smaller the frequency changes become for fusion and separation. (cf. Nabelek et al., 1970: p.545, The Journal of the Acoustical Society of America, Vol. 48). This extent for fusion or separation also depends on the geometric mean of the frequency change, as the mean extents are larger for 2000Hz than for 715Hz (cf. Nabelek et al., 1970: p.545, The Journal of the Acoustical Society of America, Vol. 48).

2.2 Nonlinear Glides

The aforementioned investigations are based on linear frequency glides. Now it is important to have a look at the discrimination of nonlinear glides as these nonlinear glides play an important role in this bachelor thesis.
Therefore the paper *Discrimination of nonlinear frequency glides* from Nick Thyer and Coug Mahar (2006) is regarded, in which nonlinear glides are compared. One half of the comparison is a nonlinear glide whose instantaneous frequency differs at the temporal center from the center frequency f_c of a linear glide by $+S$ (upper glide) and the second half is a nonlinear glide whose instantaneous frequency differs at the same time from f_c by $-S$ (lower glide). The instantaneous frequency of the upper glide respectively the lower glide at this temporal center is therefore $f_{c,nonlin,up}=f_c+S$ respectively $f_{c,nonlin,lo}=f_c-S$. Duration, start frequency and end frequency of the nonlinear upper and lower glide are identical to those of the linear glide.
In this research rising and falling nonlinear glides of two different durations (50ms and 400ms) and different transition spans were investigated to find an adequate value for *S,* for which discrimination between the upper and lower

glide is possible. The results are presented in the figure below which compares on the one hand thresholds for up-glides and down-glides for certain durations and on the other hand the dependency for thresholds on direction.
It was found out that glide discrimination was possible if not only duration but also initial and final frequencies were identical (cf. Thyer and Mahar, 2006: p.2929, The Journal of the Acoustical Society of America, Vol. 119). Apart from that thresholds increase the higher the rate of transition span gets (cf. Thyer and Mahar, 2006: p.2931, The Journal of the Acoustical Society of America, Vol. 119).
As detection is based on the excitation pattern model, detection was difficult for signals with rapid frequency changes that exceed the width of an auditory filter and are of short duration (cf. Thyer and Mahar, 2006: p.2929, The Journal of the Acoustical Society of America, Vol. 119).

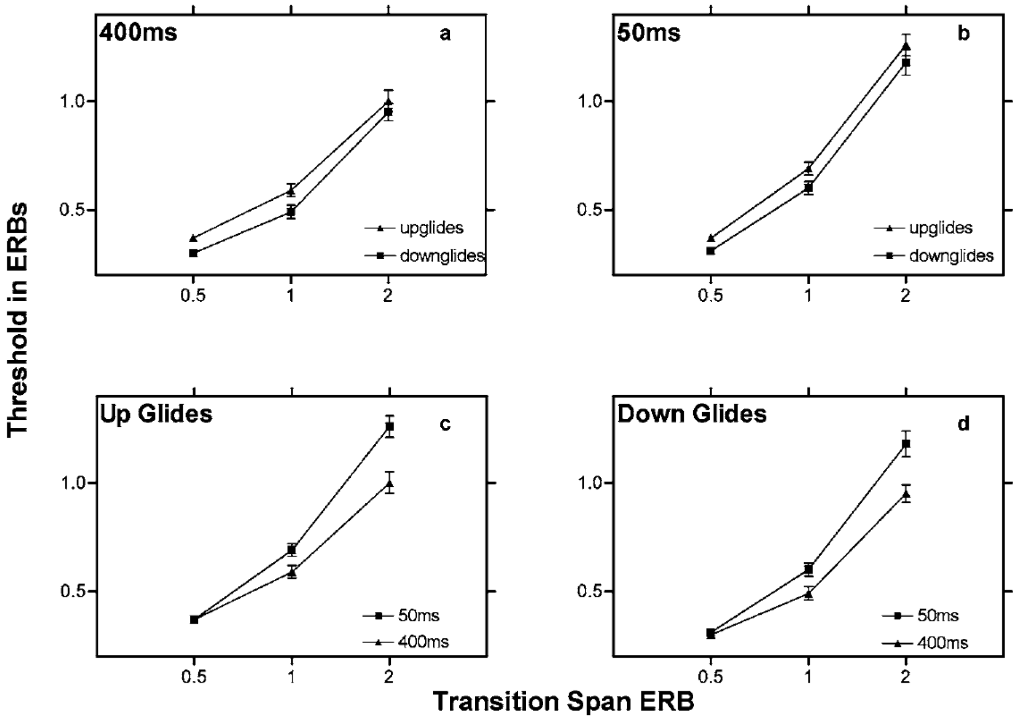

figure 9: Panels (a) and (b) results comparing mean thresholds expressed in ERBs for 400-and 50-ms glides as a function of transition span. Panels (c) and (d) results comparing mean thresholds expressed in ERBs for down glides and up glides as a function of transition span. Thyer/ Mahar (2006)

This summary only mentions some parts that have already been investigated. Of course much more research has been done on this topic but bringing up all papers that deal with glides in frequency would go beyond the scope.
Additionally the experiments mentioned above should give the reader just a short review of the topic.

3 Listening Experiment

To get an adequate form factor for which the up-glide and the down-glide evoke the same pitch, a two-alternative, forced-choice pair comparison experiment would be a convenient method. In this experiment up-glides with fixed start and end frequency, fixed arithmetic center frequency and fixed duration but with a variable form factor are compared to down-glides with same settings. It is now the main objective to change the parameter of the form factor in this way, to perceive the same pitch for the up-glide and the down-glide.
According to the results of the experiment, a psychometric function will be created which yields the point of subjective equality at which up-glide and down-glide cause the same pitch.

But before finding adequate form factors, the conditions that should be investigated in this experiment have to be chosen. After comparing existing literature with similar listening experiments (Nabelek et al., *Pitch of Tone Bursts of Changing Frequency*, 1970), a total of six conditions with the following settings were taken:

condition	bandwidth B [Hz]	arithmetic center frequency f_{ma} [Hz]	duration T [ms]
C1	50	715	40
C2	50	715	120
C3	200	715	40
C4	200	715	120
C5	200	2000	40
C6	200	2000	120

table 1: Tested conditions differing in bandwidth, arithmetic center frequency and duration

The bandwidth means the difference between the highest frequency f_{high} and the lowest frequency f_{low} the glide passes through, the arithmetic center frequency describes the arithmetic mean of f_{high} and f_{low} and the duration is the total duration of the glide that ranges from $\frac{-T}{2}$ to $\frac{T}{2}$.

3.1 Creation of Glides

With the help of *Matlab* (cf. Mathworks: https://de.mathworks.com/products/matlab.html) the function *createglide* is written which is able to create glides based on the preset parameters bandwidth *B* (in [Hz]), arithmetic center frequency f_{ma} (in [Hz]), duration *T* (in [s]), form factor *b*, direction *k* and sampling rate *Fs* (in [Hz]).

```
function[y] = createglide(B, fma, T, b, k, Fs)
% general calculations
Ts = 1/Fs;
N = floor(T * Fs);  % number of samples
ts = (-(N-1)/2):((N-1)/2); % time in samples
t = Ts * ts; % time in seconds
% lowest and highest frequency of the glide
f_low = fma – B/2;
f_high = fma + B/2;
% constant factor A
A = B/(2*sinh(b/(T/2)*T/2));
% center frequency for up-glide and down-glide
f0_up = f_low – A*(exp(-b/(T/2)*T/2) -1);
f0_down = f_high + A*(exp(-b/(T/2)*T/2)-1);
% instantaneous frequency of up-glide and down-glide
fi_exp_up = f0_up + A*(exp(b/(T/2)*t)-1);
fi_exp_down = f0_down – A*(exp(b/T/2)*t)-1;
% instantaneous phase of up-glide and down-glide
ph_exp_up = 2*pi*(f0_up * t + A*((exp(b/(T/2)*t))/(b/(T/2)) - t));
ph_exp_down = 2*pi*(f0_down * t - A*((exp(b/(T/2)*t))/(b/(T/2)) – t));
% up-glide and down-glide with amplitude of 0.8
x_exp_up = 0.8*cos(ph_exp_up);
x_exp_down = 0.8 *cos(ph_exp_down);
```

To avoid clipping and disturbing noises only 80% of the amplitude are taken.
In addition to that a suitable envelope for the created glides is needed to avoid click noises that could affect the result. For both durations (40ms and 120ms) a rise and decay time of 10ms in form of a Hann window is chosen.

```
% Hann window
W = 2* round(0.010* Fs); % 10 milliseconds rise and decay time
w = hann(W); % rise and decay in form of an Hann window
S = N - W; % area of full amplitude
e = [w(1:round(W/2));ones(S,1);w(round(W/2)+1:end)]; %glide’s envelope
```

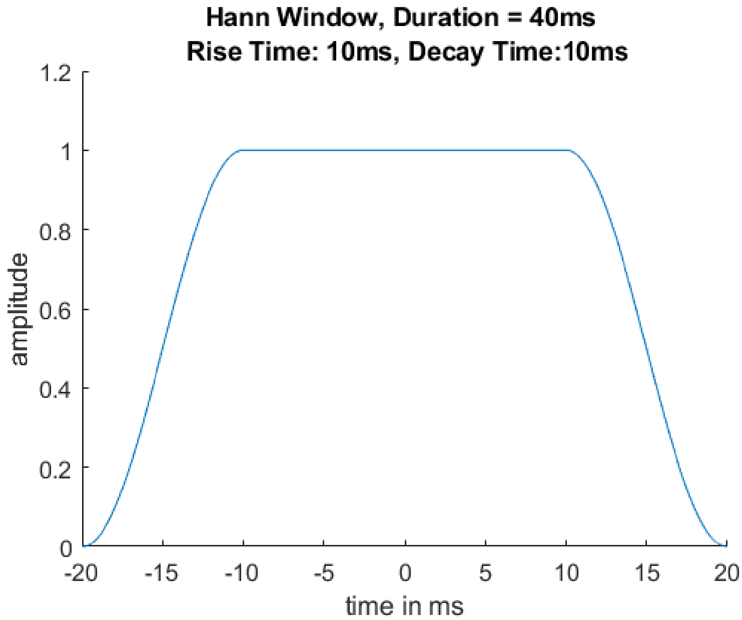

figure 10: Hann window with a duration of 40ms including rise time and decay time

Multiplying the created glides with this envelope gives the composition of the actual glides which are depicted in the picture below.

```
% glides combined with envelope
x_exp_up_hann = x_exp_up'.* e;
x_exp_down_hann = x_exp_down'.*e;
```

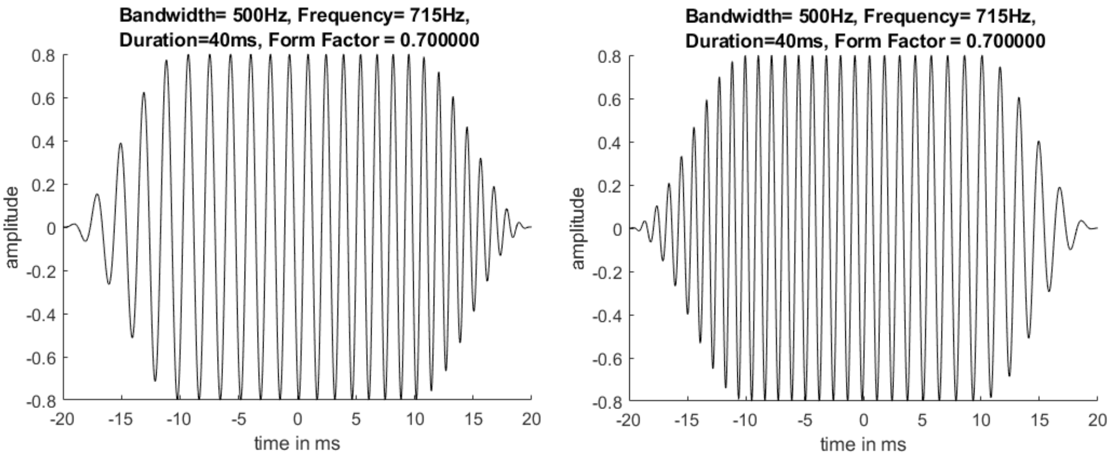

figure 11: Combination of Hann window and exponential up-glide (left) or down-glide (right)

3.2 Structure of the Experiment

According to the mentioned function *createglide*, various stimuli with different form factors are created for each condition and compared to find provisional α values that could be used for the pretest. A limitation of six α values per condition is made to prevent the listening experiment from being too long.

To find now six suitable α values for each condition, the first question to be answered was if those α values should be equally distributed or if the distance between those form factors should be variable. After short consideration it turned out to be most advantageous to take two extreme values at which the distinction between the up-glide and the down-glide is not very difficult and to make smaller intervals between the α values which lead to similar pitches for up-glides and down-glides.

α_1 is chosen at which the up-glide was clearly perceived to be higher in pitch, α_6 should evoke a higher pitch for the down-glide however. α_2 and α_5 should not yield a clear distinction in pitch but it should also not be impossible to differ between the up-glide and the down-glide. The α_2 should lead to a tendency to perceive the up-glide to be higher and for α_5 there should be a tendency to perceive a higher pitch for the down-glide. The remaining two α

values should not allow a distinction in pitch between the up-glide and the down-glide.
Those six α values for each condition are split up in to three groups, depending on the difficulty of the distinction between up-glide and down-glide. α_1 and α_6 are in the first block (easy level), α_2 and α_5 are in the second block (mid level) and α_3 and α_4 are in the third group (difficult level). Therefore each block consists of two α values.
As each α is tested for both combinations, one time the up-glide is played first and one time the down-glide is played first, two different sound files have to be created for each α .
The first sound file presents three times the up-glide with a short break of 220ms in between. After a break of 720ms the down-glide follows which is presented again three times with a distance of 220ms in between.
The second sound file has the same structure but presents the down-glide first.
Within each block each of the four different sound files is presented five times which leads to a total number of ten comparisons per α and to a number of twenty comparisons per block.

```
x_exp_up_hann_pause = vertcat(x_exp_up_hann, zeros(10000,1));
x_exp_down_hann_pause = vertcat(x_exp_down_hann, zeros(10000,1));
% combining 3 glides with a break of 220ms in between
y1 =[x_exp_up_hann_pause;x_exp_up_hann_pause;x_exp_up_hann_pause];
y2 =
[x_exp_down_hann_pause;x_exp_down_hann_pause;x_exp_down_hann_pause];
% either up-glide first or down-glide first; break of 500ms
if k==1
    y = vertcat(y1, zeros(round(44100/2),1), y2); % up down
elseif k==2
    y = vertcat(y2, zeros(round(44100/2),1), y1); % down up
end
```

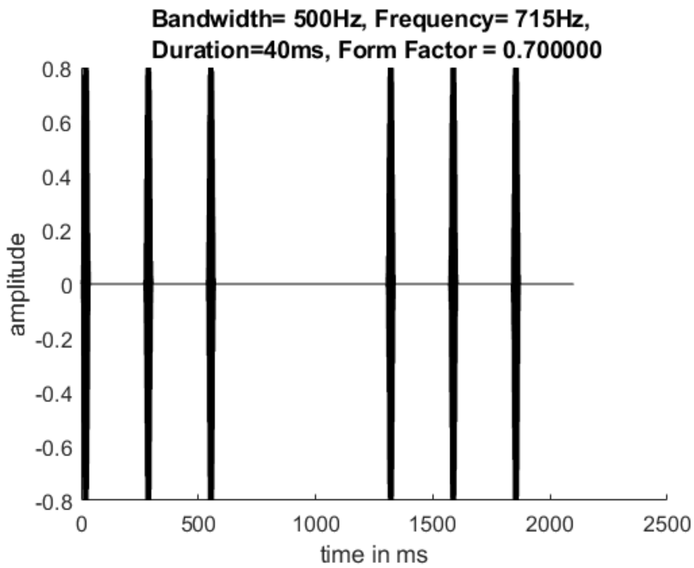

figure 12: Structure of a sound file consisting of 3 times up-glide followed by 3 times down-glide

3.3 Software

The connection between the created sound files and the graphical interface for the listening experiment is based on *Matlab* and *Pure Data* (cf. Pure Data: https://puredata.info/). As it is not really advantageous to present the same sound file and therefore the same comparison five times in a row but to mix all twenty comparisons within each block, *Matlab* creates a text file for each block which contains a random order in which the twenty comparisons are played. This order is different for each block, each condition and each subject.

Apart from that the condition order changes for each subject too. As the experiment takes about 30 minutes for each subject, the fatigue effect of the subjects must be considered which would yield lower concentration for the last tested conditions. Therefore it is important to start and end with different conditions for each subject, so that every condition is tested in times of higher and lower concentration. This is again realised by creating text files in *Matlab* which contain different condition orders in which the conditions are played for each subject.

Pure Data uses this text files to present the sound files to the subjects according to the order in the text files mentioned above.
Subjects have now the possibility to listen to the presented sound file as often as needed and to answer by clicking into the left or the right circle. Clicking into the left circle means that the first stimulus was higher in pitch and the right circle means that the second stimulus was higher. The 'play'-button is used to play the sound file again before answering.

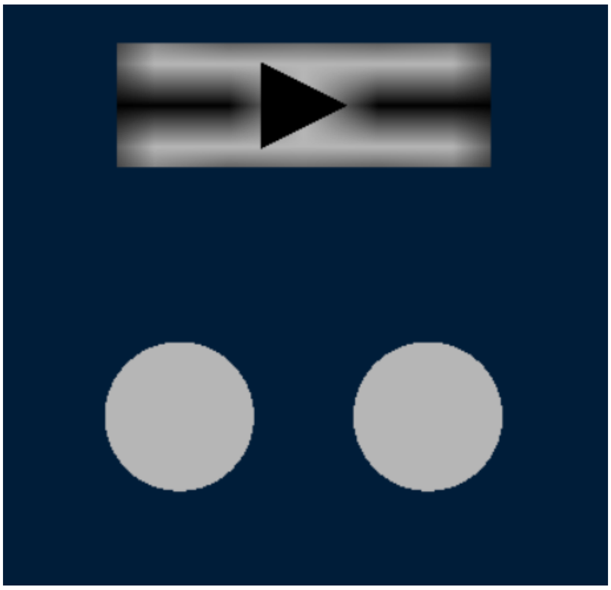

figure 13: Graphical design of the listening experiment

Each answer is then stored in a separate text file which contains not only the answer but also the condition number and the stimulus number. *Matlab* reads this text files in and creates then a psychometric function for each condition according to the answers for each α . This resulting psychometric function is a function of the tested form factors and shows the percentage for which the down-glide was perceived as higher in pitch than the up-glide.

3.4 First Pretest

According to the distinction clues mentioned above which should be considered while choosing suitable form factors, the form factors for the first pretest are determined and presented in the following table:

condition	α_1	α_2	α_3	α_4	α_5	α_6
C1	-1.0	-0.2	0.2	0.4	0.8	1.6
C2	-1.0	-0.2	0.2	0.4	0.8	1.6
C3	-1.0	0.0	0.4	0.6	1.0	1.6
C4	-1.0	0.0	0.4	0.6	1.0	1.6
C5	-1.0	0.0	0.4	0.6	1.0	1.6
C6	-1.0	0.0	0.4	0.6	1.0	1.6

table 2: Form factors for the first pretest

For the first two conditions three persons took part in the pretest and for the remaining four conditions two subjects were tested. The results are presented below in form of a psychometric function. On the abscissa the tested form factors are presented and on the ordinate the percentage of answers for which the down-glide was perceived as higher in pitch than the up-glide is shown.

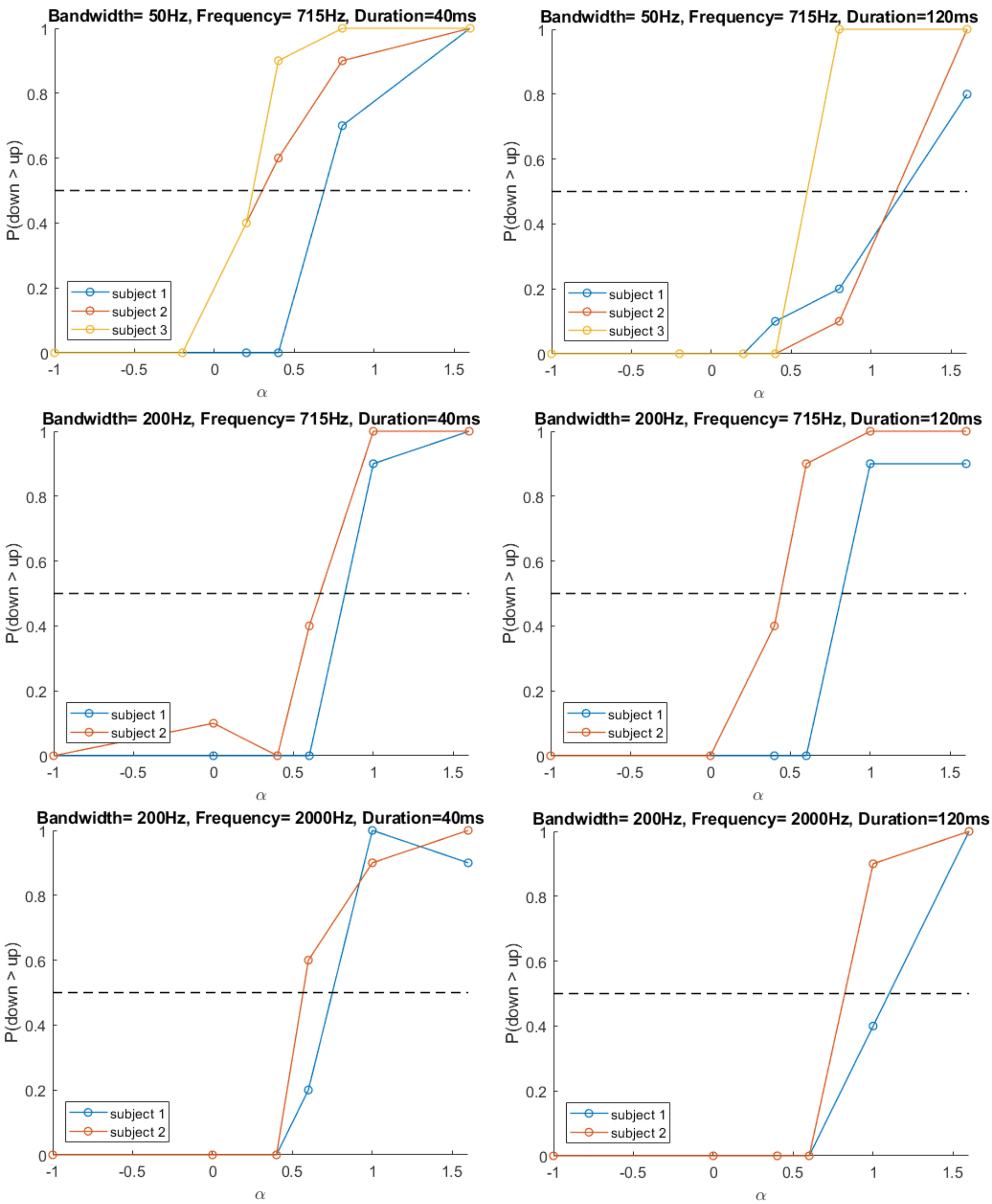

figure 14: Result of the first pretest (psychometric function)

Unfortunately the results are not satisfying because of the hard transition from 0% P(down>up) to 100% due to few values in between.
On the other hand subjects are listening in patterns in this test mode. This happens if the same stimuli are compared too often without any change. As a result, subjects keep the previous comparisons in mind while listening to the

current trials and are therefore able to identify the higher stimulus only after listening to the first stimulus. This causes the problem that people are not comparing two stimuli anymore if they have compared the same two stimuli some trials before. Subjects just answer the same as in previous trials and do not change their mind which could also be a reason for the hard transition from 0% to 100%.

3.5 Second Pretest

To avoid hearing in patterns it is important that all comparisons are different. This can be achieved by using frequency-roving. Frequency-roving means that the arithmetic center frequency f_{ma} changes for each comparison by adding or subtracting a multiple of 0.9% of the original arithmetic center frequency.
Apart from that the idea with different blocks for different distinction categories is rejected and only five instead of six α values are going to be tested for each condition. This means that there is now only one block per condition which consists of five different form factors.
Again various stimuli are created by using the function *createglide* and compared to each other to find adequate form factors for the second pretest. α_1 and α_5 should allow an easy distinction between the up-glide and the down-glide whereas the distinction based on the other three α values should be more difficult and ideally not clear.
The final α values for the second pretest are presented below.

condition	α_1	α_2	α_3	α_4	α_5
C1	-0.1	0.2	0.3	0.5	1.1
C2	-0.1	0.2	0.4	0.6	1.0
C3	0.0	0.3	0.5	0.6	0.9
C4	0.0	0.3	0.4	0.5	0.9
C5	0.2	0.5	0.6	0.7	1.0
C6	0.4	0.7	0.8	0.9	1.1

table 3: Form factors for the second pretest

Therefore each block consists of fifty different pair comparisons as each α is tested ten times but for each trial the arithmetic center frequency is different and half of the ten trials the up-glide is presented first and the other half the down-glide is first.

The following table should give an idea how the ten different stimuli at the arithmetic center frequency $f_{ma}=715\,Hz$ are created for one α considering the frequency-roving:

1. calculation of the arithmetic center frequency at *n=0*:

$$f_{ma,0}=f_{ma}-0.009\frac{f_{ma}}{2}\approx 712\,Hz$$

2. calculation of the remaining arithmetic center frequencies based on $f_{ma,0}$:

$$f_{ma,n}=f_{ma,0}+n\,0.009\,f_{ma,0}$$

stimulus	1	2	3	4	5	6	7	8	9	10
n	-4	-3	-2	-1	0	1	2	3	4	5
f_{ma} in [Hz]	686	693	699	705	712	718	725	731	737	744
order	du	ud	du	ud	du	ud	du	ud	du	ud

table 4: Calculation of the arithmetic center frequencies for the frequency-roving

As mentioned above each arithmetic center frequency is built by adding or subtracting a multiple of 0.9% of $f_{ma,0}$ to the arithmetic center frequency at *n=0*. ‘du’ means that the down-glide is presented first followed by the up-glide and ‘ud’ means that the up-glide is played before the down-glide.

Considering the α values and the frequency-roving mentioned above, a total number of fifty sound files per block is created and a second pretest with three runs is carried out.

By using the toolbox *Palamedes* (cf. Kingdom and Prins: Psychophysics. A Practical Introduction) in *Matlab*, a psychometric function averaged over all three runs is created to show the results of the second pretest.

According to this psychometric function, five α values for each condition are determined which lead to equally distributed percentage values P(down>up) between 5% and 95% at which the down-glide evokes a higher pitch than the up-glide.

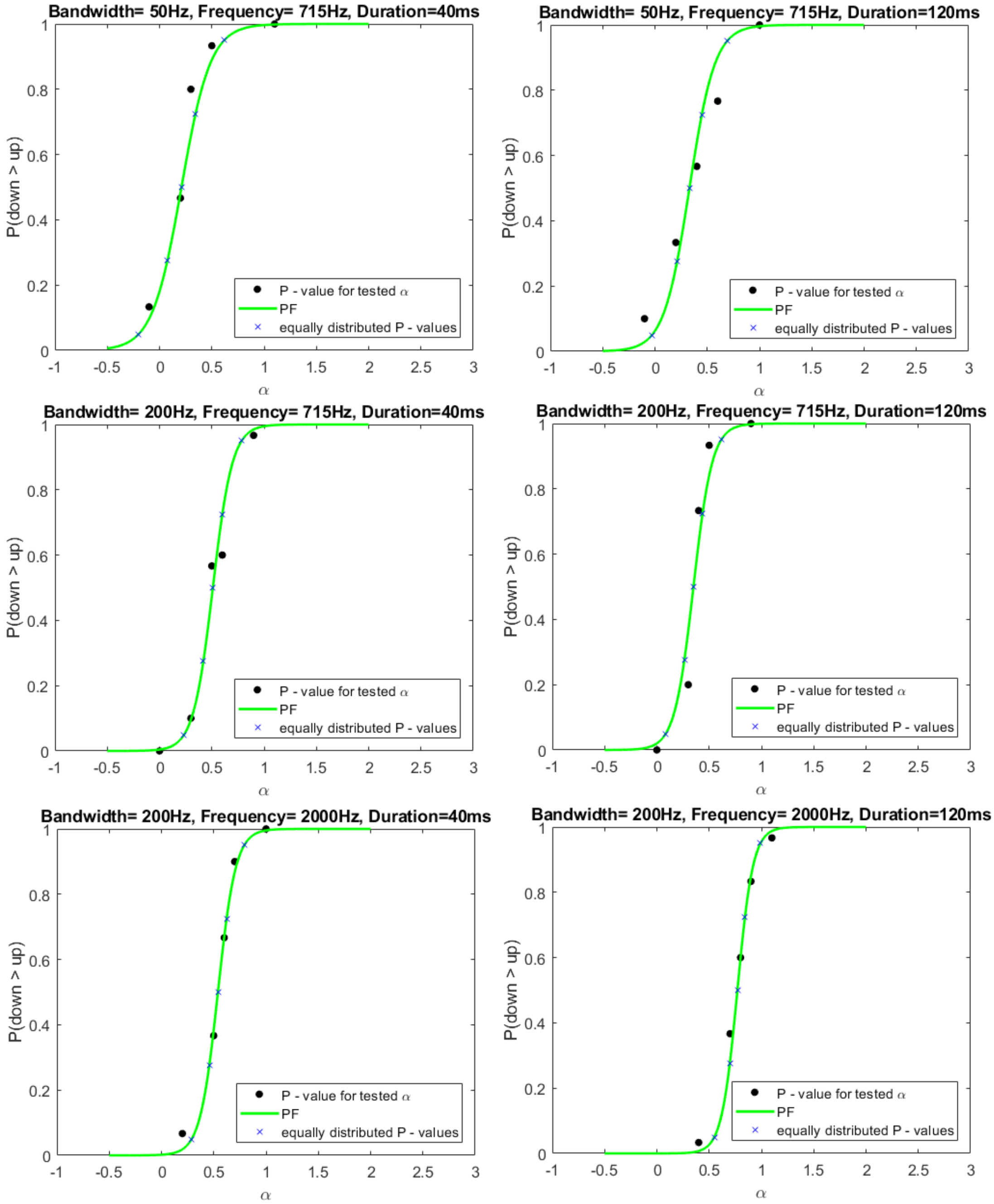

figure 15: Result of the second pretest (psychometric function) with equally distributed percentage values from 5% to 95%

condition	α_1 (5%)	α_2 (27.5%)	α_3 (50%)	α_4 (72.5%)	α_5 (95%)
C1	-0.199	0.075	0.210	0.345	0.619
C2	-0.030	0.212	0.330	0.448	0.689
C3	0.235	0.419	0.510	0.601	0.785
C4	0.085	0.263	0.350	0.437	0.615
C5	0.288	0.457	0.540	0.623	0.792
C6	0.549	0.697	0.770	0.843	0.990

table 5: Form factors for the first part of the experiment

Based on these form factors, the sound files for the actual listening experiment are created by using the function *createglide*.

3.6 First Part of the Listening Experiment

The first listening experiment took place on the 12th and 13th of April in the Inffeldgasse 10, room IG1001001. A total of 11 people participated in the experiment.

Each α in the table above (*table 5*) was tested ten times, each time with a different arithmetic center frequency to avoid hearing in patterns and half of it the up-glide was presented first and half of it the down-glide came first.

Again the sound files were created according to the pattern described above and presented in the order of the written text files. Apart from that the condition order was mixed too to avoid too high influence of fatigue effects.

Each of the six conditions included fifty pair comparisons which were presented to the subjects.

It was now on the subjects to decide for each sound file whether the first or the second stimulus was higher in pitch. Depending on the perception, subjects chose the first or the second option.

The results were stored in a text file to help with the analysis afterwards.

Mean and standard deviation of the result are presented in the following illustration.

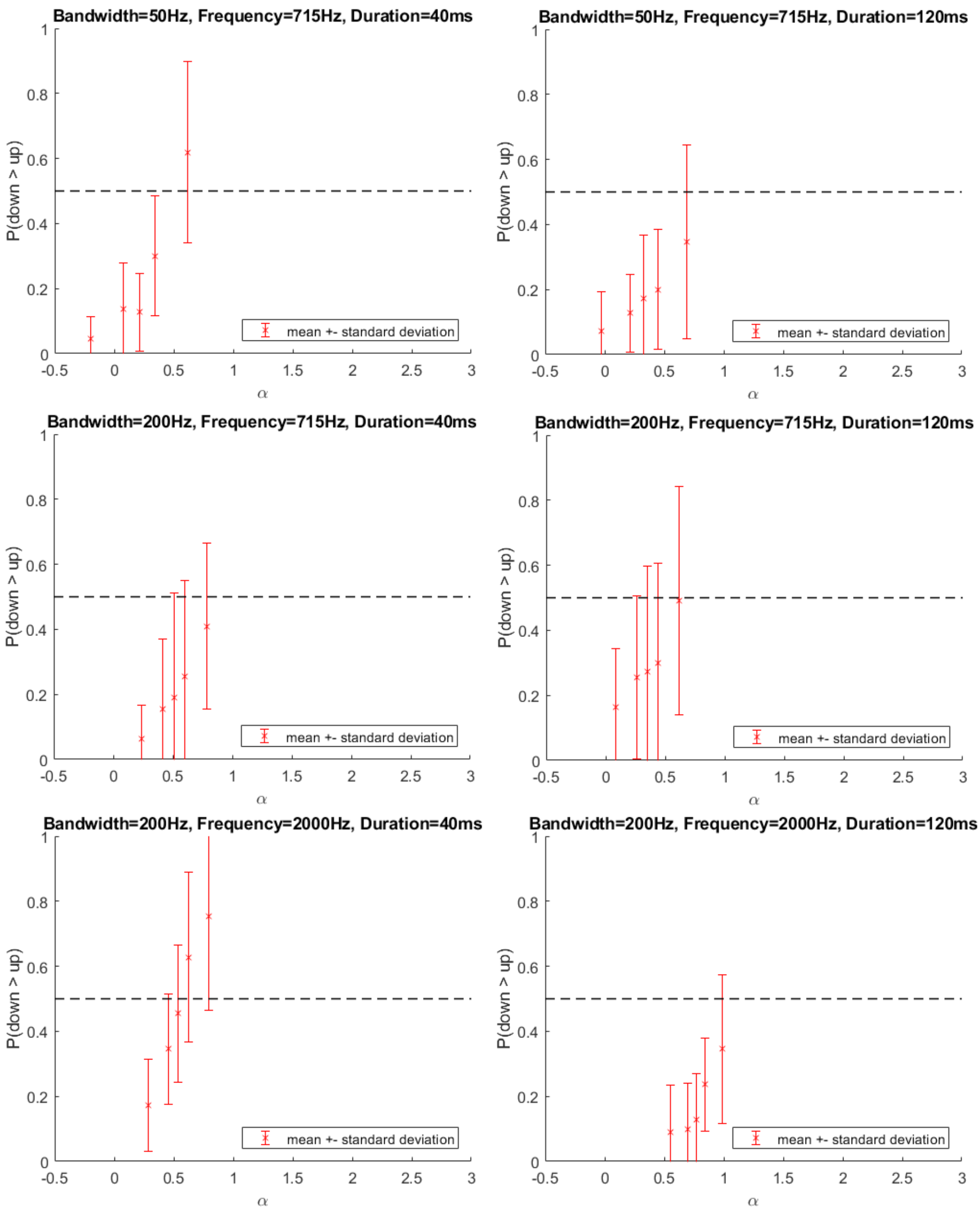

figure 16: Result of the first listening experiment (mean and standard deviation)

As the picture shows the result of the first experiment was not really satisfying. Only in two out of six conditions the percentage P(down>up) crosses the 50%-threshold. The results of conditions 2, 3, 4 and 6 do not provide any useful outcome as only the intersection of the psychometric function with the 50%-threshold would give the wanted point of subjective equality. But apparently in this experiment mainly those form factors were tested for which the up-glide evokes a higher pitch than the down-glide. Therefore the area of form factors which leads to a higher pitch for down-glides than for up-glides has to be investigated.

As can be seen in *figure 16*, the last tested form factor α_5 is mainly located near the 50%-threshold. Hence it would be useful to take this value as the lowest form factor for the new listening experiment. This also has the advantage that this value will be then tested twice which leads to a higher accuracy for this form factor near the 50%-threshold.

Based on this form factor, two higher form factors have to be found for each condition so that the down-glide causes a higher pitch than the up-glide. After creating and analysing glides with higher form factors, the following form factors are chosen for the second part of the listening experiment:

condition	α_5	α_6	α_7
C1	0.619	0.800	1.500
C2	0.689	1.000	1.500
C3	0.785	1.200	2.000
C4	0.615	1.000	1.500
C5	0.792	1.000	1.300
C6	0.990	1.800	2.500

table 6: Form factors for the second part of the experiment

As can be seen in the table above, the first form factor of the second part is identical to the last form factor of the first part of this experiment.

3.7 Second Part of the Listening Experiment

Considering the two additional form factors, the second part of the listening experiment took place on the 19^{th} and 20^{th} of April at the same place and with the same number of subjects as in the first part. The test mode stayed the same as well. Again six conditions were tested but this time only three α values per condition which leads to a number of thirty trials per condition instead of fifty.

Both, the order of the conditions and the trial order were mixed again for each subject.

Combining the results of the two parts leads to the following outcome:

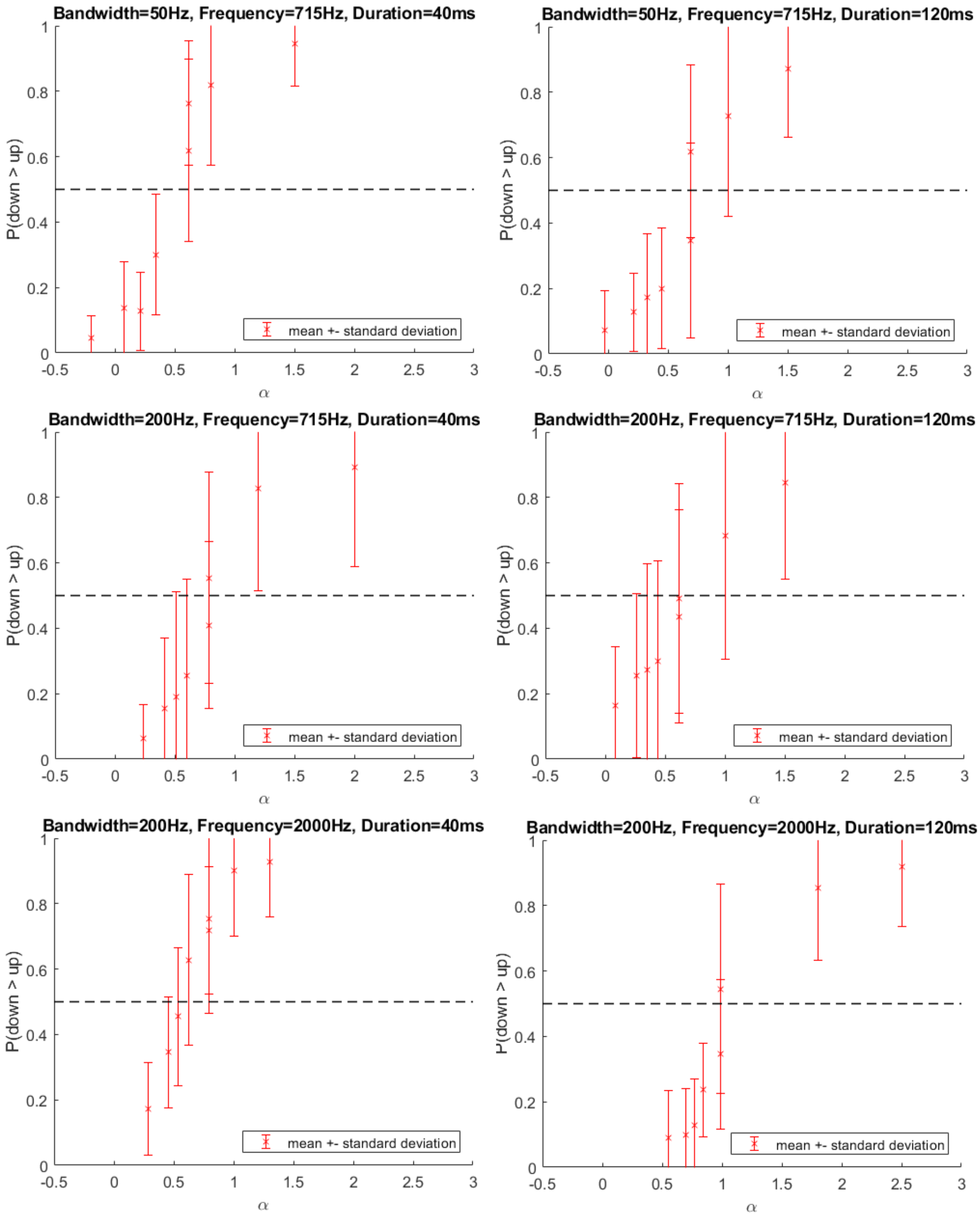

figure 17: Combination of the results of the two listening experiments (mean and standard deviation)

This figure shows that the form factor is not a linear quantity.

3.8 Final Result

According to the plot with mean and standard deviation, the two combined parts yield an acceptable result. Finally the results are analysed by the *Palamedes* toolbox and the psychometric function is fitted into the data.

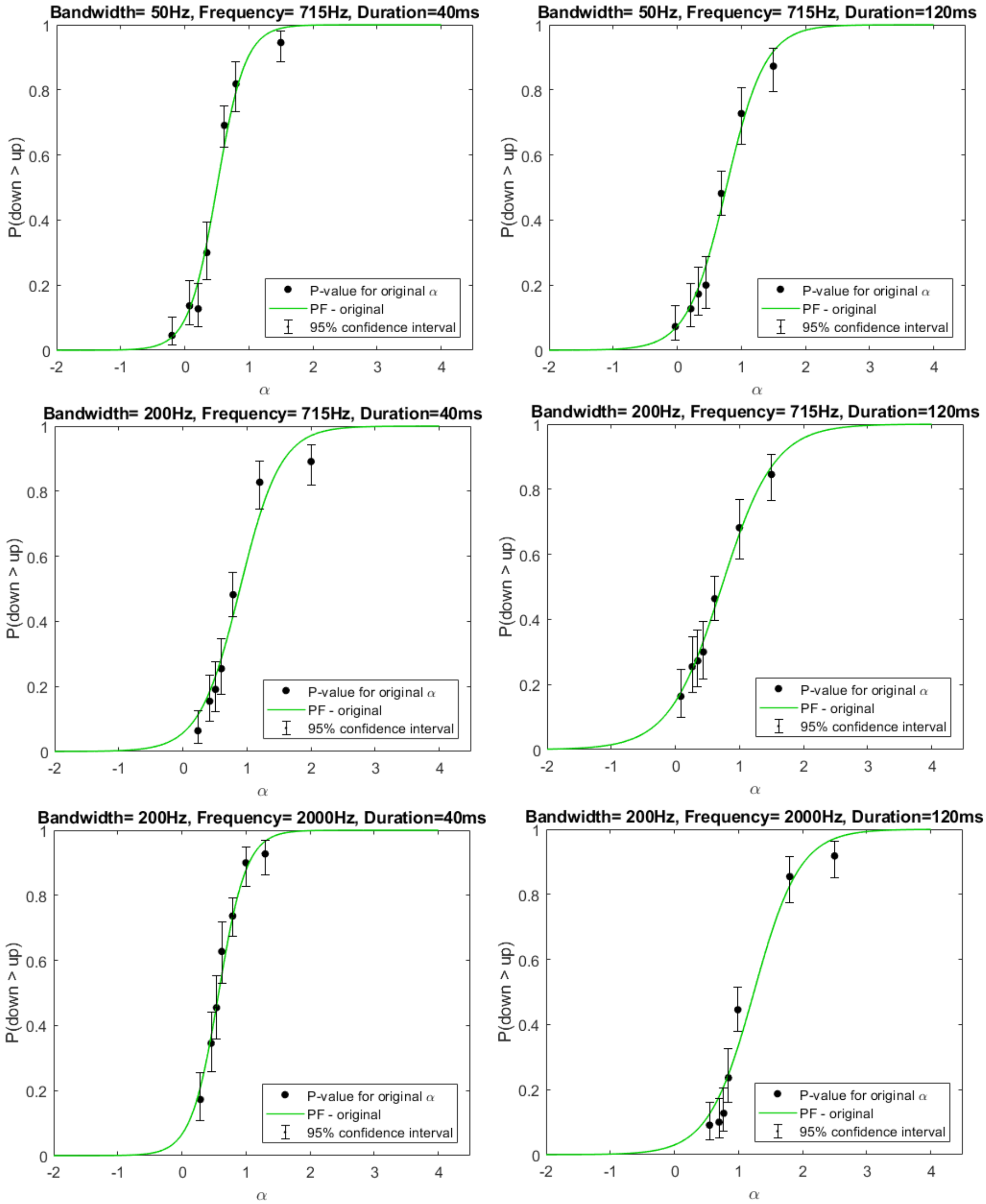

figure 18: Combination of the results of the two listening experiments (psychometric function)

The figure above shows the 95% confidence interval for each tested form factor. These confidence intervals indicate the uncertainty of the obtained percentage values and are described in more detail in the chapter *Statistical Analysis*.

To answer the main question, one could look at the psychometric function and determine the form factor α at which the psychometric function crosses the threshold of 50% P(down>up).
But on closer examination it seems that the psychometric function does not fit the data points very well. To investigate how well the psychometric function fits the data, the *Palamedes* toolbox offers a function to determine a *goodness-of-fit* value. This value is a value between 0 and 1 and the higher this value gets, the better the fitting is. A *goodness-of-fit* value smaller than 0.05 is unacceptable (cf. Kingdom and Prins, 2010: p.73).
Applying this *goodness-of-fit* function to the current psychometric function shows that the *goodness-of-fit* value is in some conditions acceptable but not really satisfying.

condition	C1	C2	C3	C4	C5	C6
goodness-of-fit	0.00	0.15	0.00	0.80	0.19	0.00

table 7: Goodness-of-fit values of the original psychometric function

One possibility to achieve higher *goodness-of-fit* values is to transform the α values before fitting the psychometric function. As it seems that the form factor is not a linear quantity (*figure 17*), one could transform the ordinate by using for example a logarithmic function or an exponential function.
Various combinations are tested, also a shifting of the α values along the ordinate before applying the logarithm or the exponential function. For various combinations the *goodness-of-fit* values are calculated and compared to each other. It becomes obvious that there exists no uniform transformation that fits best for all conditions. For each condition an individual transformation has to be

made to achieve higher *goodness-of-fit* values. These individual transformations and the resulting *goodness-of-fit* values are presented below.
According to the values in the following table (*table 8*) a transformation is made using the following equation:

$$\alpha_{transf} = (\alpha_{original} + v)^h$$

condition	shifting (v)	exponent (h)	*goodness-of-fit*
C1	0.9	0.2	0.02
C2	0.9	0.6	0.33
C3	0.1	0.1	0.11
C4	0.7	0.5	0.95
C5	0.1	0.3	0.67
C6	-0.1	0.05	0.05

table 8: Transformation of the original form factors with the resulting goodness-of-fit values

Based on these transformed α values, again a psychometric function is fitted into the data. Afterwards this obtained psychometric function is back transformed into the original domain by back transforming α_{transf} into $\alpha_{original}$.
The following graphic shows that this back transformed psychometric function leads to a better fitted function for the α values than the original function.

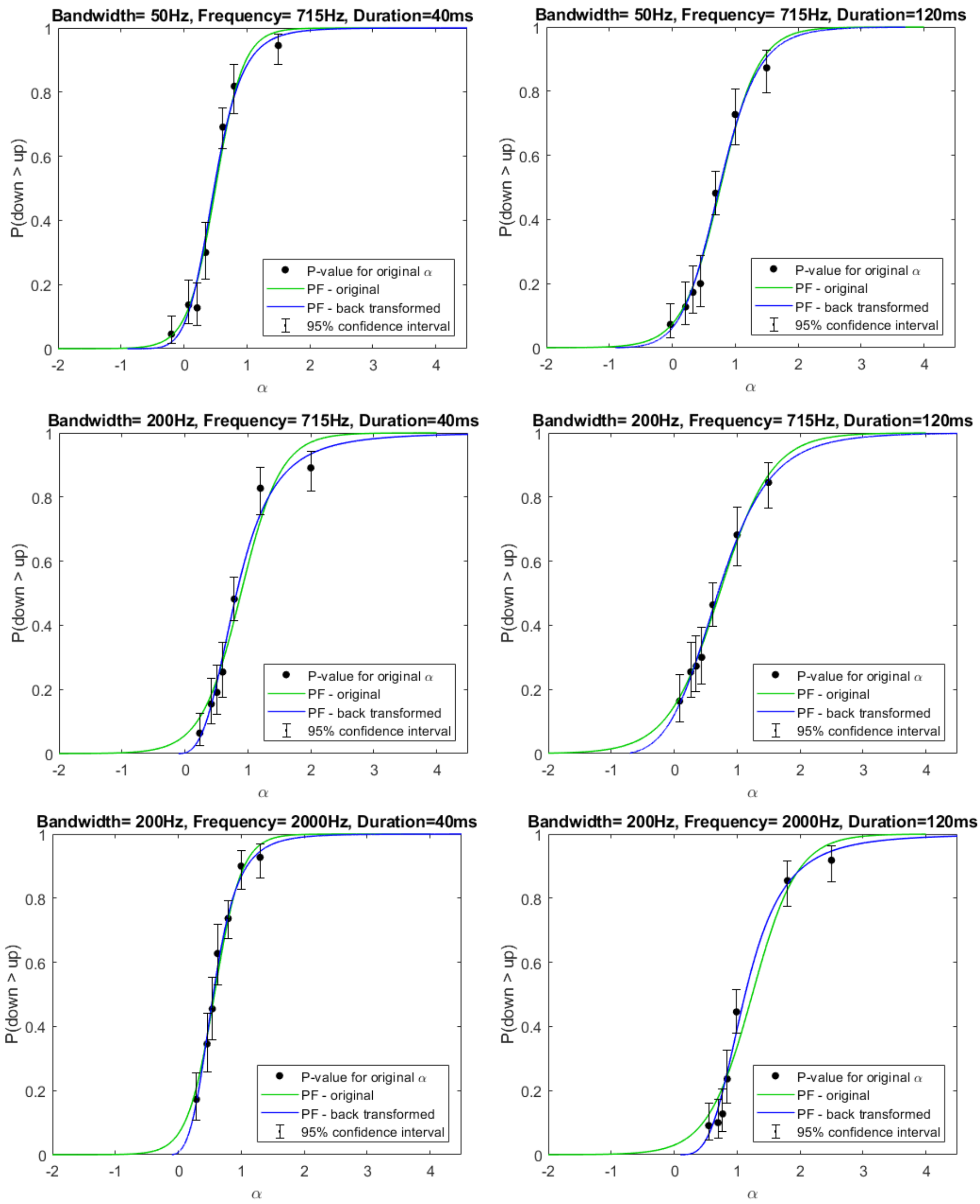

figure 19: Comparison between the original psychometric function and the back transformed psychometric function

Now the demanded form factor can be obtained from this psychometric function. There are two possibilities to get the $\alpha_{50\%}$. One way is to have a look at the psychometric function and read out the form factor at which the psychometric function crosses the 50%-threshold.

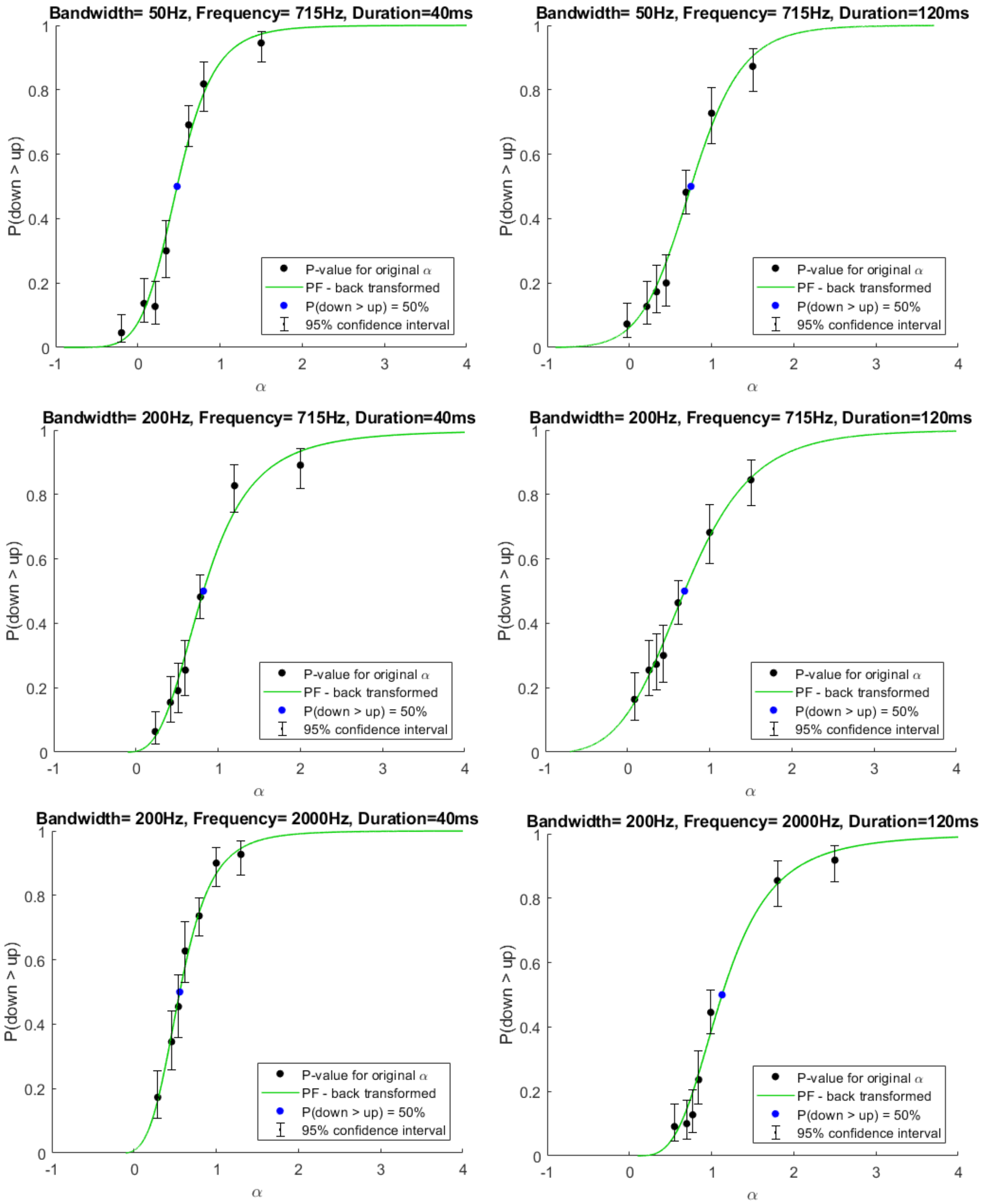

figure 20: Form factor at the 50%-threshold for the back transformed psychometric function

But as this way is not really exact, a second option is taken into account. The form factor at which the psychometric function crosses the 50%-threshold is interesting. Therefore the functional equation of the psychometric function is determined and is then set to 0.5.

The transformed function corresponds to the logistic function with the following functional equation (cf. Kingdom and Prins*,* 2010: p.82):

$$f_{Logistic}(x)=\frac{1}{1+e^{-\beta(x-\alpha)}}$$

In this equation α describes the ordinate value where the logistic function has its midpoint and β determines the slope of the curve (cf. Kingdom and Prins*,* 2010: p.82).

α is the value of the form factor at the 50%-threshold. This value is given by the *Palamedes* toolbox as well as β which describes the slope at $\alpha_{50\%}$. This functional equation describes the transformed function as mentioned above. To get the functional equation of the back transformed psychometric function, *x* is replaced by the transformed *x*:

$$x_{transf}=(x+v)^h$$

This leads to the functional equation of the back transformed psychometric function:

$$f_{Logistic,backtransf}(x)=\frac{1}{1+e^{-\beta((x+v)^h-\alpha)}}$$

To get the form factor at the 50%-threshold, this functional equation is set to 0.5.

$$f_{Logistic,backtransf}(x)=\frac{1}{1+e^{-\beta((x+v)^h-\alpha)}}=0.5=\frac{1}{2}$$

$$1+e^{-\beta((x+v)^h-\alpha)}=2$$

$$e^{-\beta((x+v)^h-\alpha)}=1$$

$$-\beta((x+v)^h-\alpha)=0$$

$$((x+v)^h-\alpha)=0$$

$$(x+v)^h=\alpha$$

$$\alpha^{\frac{1}{h}}=x+v$$

$$\alpha^{\frac{1}{h}}-v=x=\alpha_{50\%,backtransf}$$

For each condition the psychometric function has a different form as the parameters α and β have different values for each condition. Therefore each condition has an individual functional equation for the psychometric function and hence a different $\alpha_{50\%,backtransf}$.

The following table gives an overview of the essential values for each condition which lead according to the equation above to the form factor at the 50%-threshold:

condition	shifter (v)	exponent (h)	α	$\alpha_{50\%,backtransf}$
C1	0.9	0.2	1.066	0.478
C2	0.9	0.6	1.351	0.750
C3	0.1	0.1	0.992	0.821
C4	0.7	0.5	1.181	0.695
C5	0.1	0.3	0.881	0.556
C6	-0.1	0.05	1.001	1.128

table 9: Essential parameters for the calculation of the form factor at the 50%-threshold and the resulting form factor at this threshold

3.9 Statistical Analysis

Now a statistical analysis based on these psychometric functions is performed. In the pictures above the percentage value P(down>up) for each tested form factor is given as the mean value over all subjects. But as only a few persons took part in the listening experiment, there is always an uncertainty of the obtained result. The more subjects are asked in the experiment, the smaller the area gets in which the result is located.

3.9.1 95% Confidence Interval

In this case it is advantageous to determine a 95% confidence interval for each tested form factor. This interval indicates the area in which the percentage value is located with a probability of 95%. This confidence interval is calculated with

the *Matlab* function *binofit* and depends on the number of answers for which the down-glide was higher than the up-glide and the total number of trials.
As mentioned above the more subjects are tested, the smaller this confidence interval gets. This fact can be seen in the graphics above. The α_5 was tested twice as often as the other form factors. Therefore the 95% confidence interval for this α value is discernible smaller than the interval for the other six form factors.

3.9.2 Horizontal Standard Deviation at the 50%-Threshold

In addition to the uncertainty of the percentages, the deviation of the obtained $\alpha_{50\%}$ is important. To get this uncertainty, the *Palamedes* toolbox plays an essential role. The *bootstrapping* function in *Palamedes* gives the standard deviation for the $\alpha_{50\%}$ and for the slope at this point (cf. Kingdom and Prins, 2010: p.72).
As the final psychometric function was not created by *Palamedes* itself, it is not possible to apply this *bootstrapping* function on this psychometric function. Therefore the standard deviation for the transformed psychometric function is calculated by using the *bootstrapping* function in *Palamedes* and then this standard deviation is back transformed into the original domain according to the equation below. The results are presented in *table 10*.

$$\alpha_{50\%,backtransf} \pm SD_{backtransf} = (\alpha_{50\%,transf} \pm SD_{transf})^{\frac{1}{h}} - v$$

condition	$\alpha_{50\%,transf} \pm SD_{transf}$	$\alpha_{50\%,backtransf} - SD_{backtransf}$	$\alpha_{50\%,backtransf} + SD_{backtransf}$
C1	1.066 ± 0.003	0.478 - 0.019	0.478 + 0.019
C2	1.351 ± 0.014	0.750 - 0.027	0.750 + 0.028
C3	0.992 ± 0.003	0.821 - 0.026	0.821 + 0.027
C4	1.181 ± 0.014	0.695 - 0.033	0.695 + 0.033
C5	0.881 ± 0.007	0.556 - 0.017	0.556 + 0.017
C6	1.001 ± 0.002	1.128 - 0.031	1.128 + 0.031

table 10: Mean and standard deviation of the form factor at the 50%-threshold

3.9.3 Slope at the 50%-Threshold

Finally the slope of the psychometric function at the form factor $\alpha_{50\%}$ is interesting, as the slope is an indicator for the precision of the obtained result and is inversely proportional to the standard deviation (cf. Kingdom and Prins, 2010: p.21). To get this slope there are basically two methods: building the analytical derivation or the numerical derivation.

First the numerical derivative is calculated as this is the easier and faster option. In addition the analytical derivative is built too to check the numerical way.

3.9.3.1 Numerical Derivation

To get the numerical derivative the sampled psychometric function is taken. As the distances on the ordinate between those sampled values are not equidistant because of the back transformation, these individual distances Δx have to be calculated first. Considering these individual distances, the difference quotient is calculated according to the following equation:

$$k=\frac{f(x_0+\Delta x)-f(x_0)}{\Delta x}$$

In this equation *k* describes the slope of the secant that runs through $f(x_0)$ and $f(x_0+\Delta x)$.

As the psychometric function consists of 10000 sampling points, this psychometric function is subdivided into 9999 intervals for which the slope *k* is calculated according to the equation above.

```
% numerical derivation
x_dist = zeros(6,9999);   % distance vector
dfnum = zeros(6,9999);    % derivative vector
for r = 1:6
      for n = 2:10000
            x_dist(r, n-1)=x_btransformed(r,n)-x_btransformed(r,n-1);
      end
end
```

```
% numerical derivation dfnum for each distance
for s = 1:6
       for l = 2:10000
         dfnum(s, l-1) = (Fit(s, l) - Fit(s, l-1))/x_dist(s, l-1);
       end
end
```

Plotting these 9999 slope values makes it possible to get the derivative of the psychometric function, which is presented below. The interesting point is the derivative at the form factor $\alpha_{50\%}$. This point is marked with a red circle.

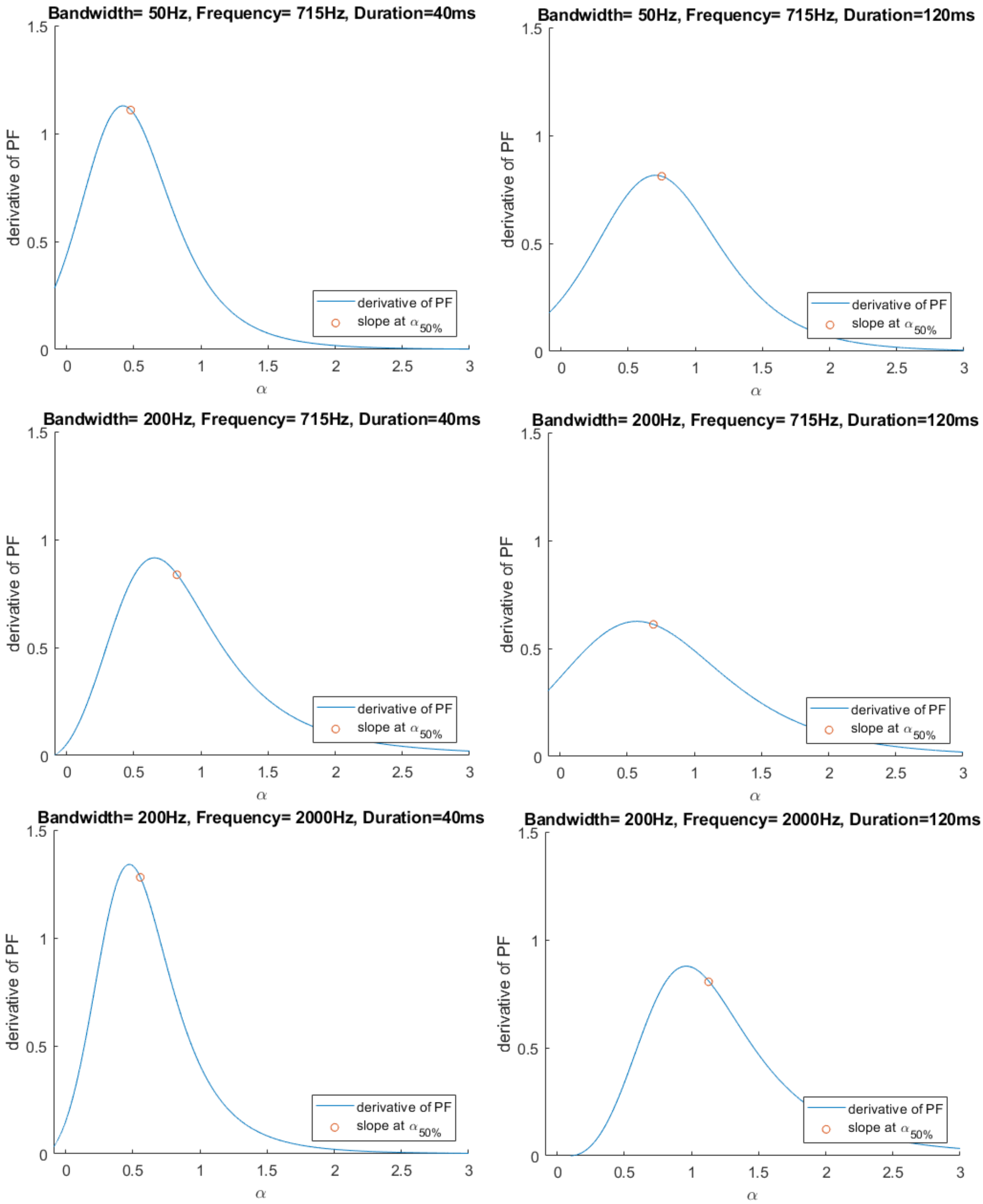

figure 21: Numerical derivative of the psychometric function

3.9.3.2 Analytical Derivation

For this option the functional equation of the back transformed psychometric function is taken into account. As mentioned above, this functional equation is described as follows:

$$f_{Logistic,\,backtransf}(x)=\frac{1}{1+e^{-\beta((x+v)^h-\alpha)}}$$

To get the analytical derivative of the psychometric function at $\alpha_{50\%}$, this functional equation is derived, afterwards α , β , *v* and *h* are inserted for each condition individually according to the table above (*table 9*) and *x* is set to $\alpha_{50\%}$.

$$\frac{d}{dx}f_{Logistic,\,backtransf}(x)=\frac{d}{dx}\left(\frac{1}{1+e^{-\beta((x+v)^h-\alpha)}}\right)=\frac{\beta h(v+x)^{h-1}e^{-\beta((v+x)^h-\alpha)}}{\left(e^{-\beta((v+x)^h-\alpha)}+1\right)^2}$$

Both, the numerical and the analytical derivation, lead to the same result. Therefore the slope at $\alpha_{50\%}$ is specified as follows:

condition	C1	C2	C3	C4	C5	C6
slope at $\alpha_{50\%}$	1.109	0.811	0.838	0.611	1.281	0.807

table 11: Slope of the psychometric function at the 50%-threshold for each condition

4 Summary

4.1 Was the Result Satisfying?

Now the most significant question arises if the listening experiment leads to a satisfying result or if some parts remain unanswered.
The particular attention of this bachelor thesis is paid on finding a form factor at which the up-glide and down-glide are perceived to be identical in pitch. Therefore the form factor at which the psychometric function crosses the 50%-threshold is the most important value.
Unfortunately in the first listening experiment this 50%-threshold was not reached by the psychometric function in all conditions so that a second part of the listening experiment with different form factors had to be carried out.
Two more form factors were tested in this second part. This time the experiment was successful. Combining the results of the two parts leads to a satisfying final outcome.

As the psychometric function fitted by the *Palamedes* toolbox was not fitting the data points very well, a suitably chosen transformation has to be found to get a better fitted psychometric function. A following back transformation of this psychometric function delivers then a psychometric function in the original domain that fits the data points very well.
With assistance of this psychometric function, the form factor at the point of subjective equality could be read out. Another way to get the $\alpha_{50\%}$ is to calculate this value by setting the functional equation of the psychometric function equal to 0.5.
Mean and standard deviation of the obtained form factor are presented for each condition in the table below:

condition	$\alpha_{50\%,backtransf} - SD_{backtransf}$	$\alpha_{50\%,backtransf} + SD_{backtransf}$
C1	0.478 - 0.019	0.478 + 0.019
C2	0.750 - 0.027	0.750 + 0.028
C3	0.821 - 0.026	0.821 + 0.027
C4	0.695 - 0.033	0.695 + 0.033
C5	0.556 - 0.017	0.556 + 0.017
C6	1.128 - 0.031	1.128 + 0.031

table 12: Form factor for each condition which leads to the same pitch for up-glide and down-glide

Out of personal interest the instantaneous frequency of up-glide and down-glide is depicted in the following plots according to the calculated form factors for each condition.

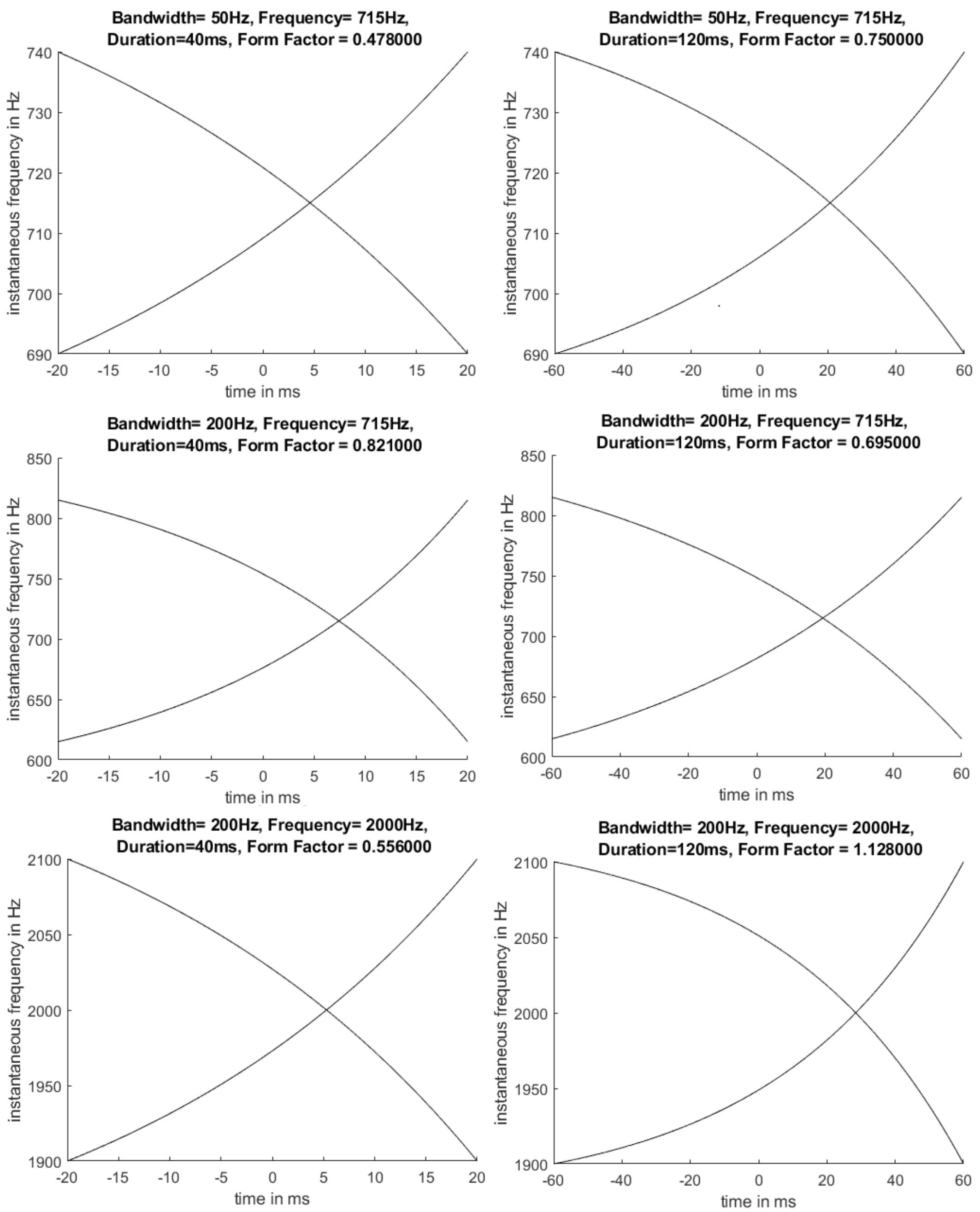

figure 22: Instantaneous frequency of up-glide and down-glide according to the calculated form factors for all conditions

This graphical representation shows that the time of the intersection is located somewhere in the rear area of the glide. The instantaneous frequency at which the intersection takes place is exactly the arithmetic center frequency.

The following table describes both the time and the frequency of intersection for each condition.

condition	frequency [Hz]	time ± SD [ms]	time in % of T ± SD [%]
C1	715	4.61 – 0.17	61.5 – 0.4
		4.61 + 0.17	61.5 + 0.4
C2	715	20.66 – 0.64	67.2 – 0.5
		20.66 + 0.64	67.2 + 0.5
C3	715	7.43 – 0.20	68.6 – 0.5
		7.43 + 0.20	68.6 + 0.5
C4	715	19.36 – 0.81	66.1 – 0.7
		19.36 + 0.79	66.1 + 0.7
C5	2000	5.30 – 0.14	63.2 – 0.4
		5.30 + 0.15	63.2 + 0.4
C6	2000	28.43 – 0.55	73.7 – 0.5
		28.43 + 0.56	73.7 + 0.5

table 13: Frequency and time of the intersection of up-glide and down-glide for each condition

What is important to mention is that the time of intersection is given in the interval of $\frac{-T}{2}$ to $\frac{T}{2}$ which means that the time described in the table above is not the time difference from the starting point but the value in this interval. For a better representation the time is given additionally as the percentage value of the whole duration *T*.

In this representation it becomes obvious that the time of intersection and therefore the time at which the perception takes place is located in the area from 60% to 75% of the whole glide duration.

A statistical analysis gives both, the 95% confidence interval of the obtained percentage values and the horizontal standard deviation at the point of subjective equality $\alpha_{50\%}$.

Finally the slope of the psychometric function is calculated by using the numerical derivation first and the analytical derivation afterwards. Both derivations amount to the same result.

4.2 Dependencies on Condition

Comparing these form factors $\alpha_{50\%}$ for all six conditions, dependencies on arithmetic center frequency, bandwidth of transition span and glide duration become clear. To get a graphical representation, the form factor $\alpha_{50\%}$ is analysed as a function of the bandwidth of the transition span and the arithmetic center frequency.

There are two different options to combine the bandwidth of the transition span and the arithmetic center frequency. On the one hand the bandwidth is expressed relatively to the arithmetic center frequency $\frac{B}{f_{ma}}$. On the other hand the bandwidth is represented as a factor of the equivalent rectangular bandwidth ERB of the arithmetic center frequency. The conversion from the arithmetic center frequency to the equivalent rectangular bandwidth is done according to

$$ERB = 24.7(4.37f + 1)$$

where f is the arithmetic center frequency in [kHz] (cf. Thyer and Mahar, 2006: p.2931, The Journal of the Acoustical Society of America, Vol. 119, cited from Glasberg and Moore (1990)).

For both durations (40ms and 120ms) the form factor $\alpha_{50\%}$ is expressed as a function of the relative bandwidth, related either to the arithmetic center frequency or to the equivalent rectangular bandwidth of the arithmetic center frequency.

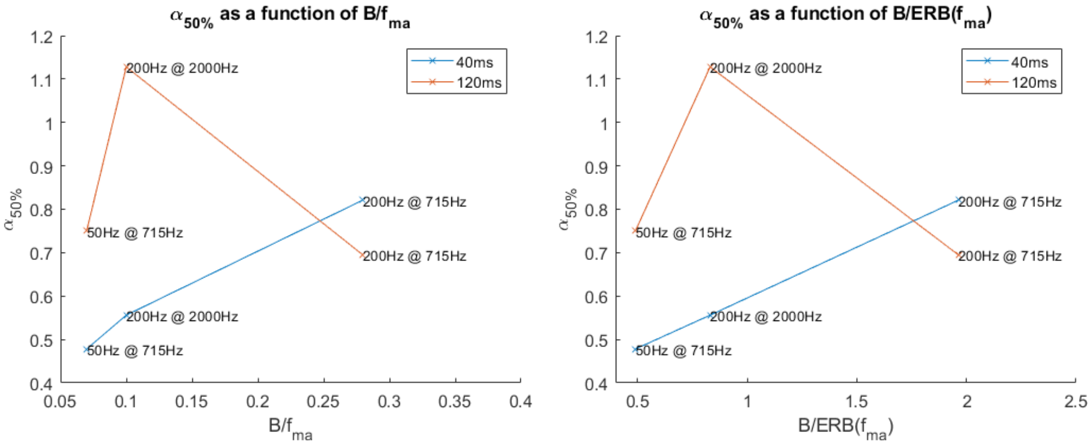

figure 23: Form factor as a function of the relative bandwidth; left: bandwidth related to the arithmetic center frequency; right: bandwidth related to the ERB of the arithmetic center frequency;

According to these plots, $\alpha_{50\%}$ is increasing linearly with increasing relative bandwidth for a duration of 40ms. For longer durations (120ms), $\alpha_{50\%}$ is increasing to a maximum and then decreasing for higher relative bandwidths. Perhaps this is because for longer durations not only one single pitch is perceived but a frequency span.

4.3 Further Research Based on this Experiment

Furthermore those results can be used for additional investigations. For instance it would be interesting to identify the perceived pitch of the glide at the determined form factor $\alpha_{50\%}$.

To find out the frequency of the perceived pitch, again a two-alternative, forced-choice pair comparison test could be used. A stimulus with changing instantaneous frequency is compared to a sine with same duration but with constant frequency. It is now on the subjects to decide whether the glide or the sine was higher in pitch. If the sine was higher in pitch, the sine's frequency decreases for the next pair comparison. If the sine is still higher in pitch the sine's frequency is decreasing again till the sine's frequency evokes a lower pitch. Then the sine's frequency increases et etcetera. Therefore the pitch of the sine comes closer to the pitch of the glide after each trial. After the procedure the frequency of the sine describes best the perceived pitch of the glide.

This knowledge of the identical pitch for up-glides and down-glides is especially important while creating discrimination contours. For the creation it is essential that glides evoke a pitch that is independent of direction.

List of References

Bibliography

Kingdom, Frederick A. A./ Prins, Nicolaas: Psychophysics. A Practical Introduction. First edition, London/ Burlington/ San Diego: Elsevier 2010.

Madden, John P. / Fire, Kevin M.: Detection and discrimination of gliding tones as a function of frequency transition and center frequency. In: The Journal of the Acoustical Society of America, Vol. 100, 1996, 3754 – 3760.

Madden, John P. / Fire, Kevin M.: Detection and discrimination of frequency glides as a function of direction, duration, frequency span, and center frequency. In: The Journal of the Acoustical Society of America, Vol. 102, 1997, 2920 – 2924.

Nabelek, Igor V./ Nabelek, Anna K./ Hirsh, Ira J.: Pitch of Tone Bursts of Changing Frequency. In: The Journal of the Acoustical Society of America, Vol. 48, 1970, 536 – 553.

Thyer, Nick/ Mahar, Dough: Discrimination of nonlinear frequency glides. In: The Journal of the Acoustical Society of America, Vol. 119, 2006, 2929 – 2936.

Internet Sources

The MathWorks, Inc: MathWorks. In: https://de.mathworks.com/products/matlab.html [04.09.2018].

Institute of Electronic Music and Acoustics Graz: Pure Data. In: https://puredata.info/ [04.09.2018].