Pitch SIPI

Experiments for Rate Pitch Perception in Cochlear-Implant Listeners with Short-Interpulse-Interval (SIPI) Stimulation

Project Report

Supported by the National Institutes of Health (NIH), grant #: R01DC005775

Martin Lindenbeck

Reviewer: Prof. Dr. Robert Höldrich
Supervisors: Doz. Dr. Bernhard Laback, Doz. Dr. Piotr Majdak

September 29, 2017

Acoustics Research Institute

Institute for Electronic Music and Acoustics

Abstract

Despite ongoing research, cochlear-implant (CI) listeners still show large deficits in pitch perception. In essence, current CIs produce amplitude modulated pulse trains as electrode signals. Here the modulation contains the entire acoustic information, the pulse trains remain periodic.

The Acoustics Research Institute (ARI) investigates a new method of stimulation: In order to increase ITD (interaural time difference) sensitivity while maintaining speech intelligibility, the temporal fine structure is modified by adding additional pulses with short-interpulse-intervals (SIPI). Ongoing experiments with pseudo-syllabic signals already showed improvements in ITD sensitivity.

Within the project, an ARI software (ExpSuite) is extended to include experiments that allow to test the hypothesis that the introduction of SIPI pulses improves the pitch perception of CI listeners. Besides, potentially confounding factors such as loudness differences are addressed.
To validate the software functionality, a pilot study is conducted. Furthermore, insight into the principal effectiveness of the SIPI pulses shall be gained to find out, what extent and which setting a subsequent, comprehensive study should have.

Zusammenfassung

Trotz andauernder Forschung haben Cochleaimplantat-TrägerInnen nach wie vor große Defizite bei der Tonhöhenwahrnehmung. Aktuelle Cochleaimplantate erzeugen im Wesentlichen amplitudenmodulierte Pulsketten als Signale für die jeweiligen Elektroden. Die gesamte akustische Information ist in der Modulation enthalten, die Pulsketten bleiben periodisch.

Am Institut für Schallforschung (ISF) wird eine neue Stimulationsart untersucht: Um bessere ITD (engl.: interaural time difference) Sensitivität bei gleichzeitiger Erhaltung der Sprachverständlichkeit zu erzielen, wird die zeitliche Feinstruktur gezielt durch Hinzufügen zusätzlicher Pulse mit kurzen Interpuls-Intervallen (engl.: short interpulse intervals, SIPI) modifiziert. Laufende Experimente mit pseudo-syllabischen Signalen zeigten bereits Verbesserungen in der ITD Wahrnehmung.

Im Rahmen dieses Toningenieur-Projekts wird eine Software des ISF (ExpSuite) um Experimente erweitert, die die Hypothese, dass die Einführung der SIPI-Pulse ebenfalls die Tonhöhenwahrnehmung von CI-TrägerInnen verbessert, testen können. Dabei wird auch auf mögliche Störfaktoren wie etwa Unterschiede in der Lautheit eingegangen.
Zur Überprüfung der Software-Funktionalität wird eine Pilot-Studie durchgeführt. Ferner sollen damit Erkenntnisse über die grundsätzliche Wirksamkeit der SIPIs gewonnen werden um herauszufinden, welchen Umfang und insbesondere welches Setting eine anschließende, vollständige Studie haben soll.

Contents

Nomenclature

Amp Signal amplitude: current of the signal without AM, i.e. MD = 0.

AM Amplitude modulation: Audio information is commonly transmitted by the CIs via AM.

CI Cochlear Implant: In general CI denotes the whole complex consisting of microphone, signal processor and the actual implant. In context with CI studies, only the implant itself is meant.

CL Comfortable level: The current at which the participant feels most comfortable to conduct the experiments.

DR Dynamic range: The difference between MCL and THR, usually in μA or dB.

ENV Slowly varying envelope of a time domain signal: It can be extracted using the Hilbert transform.

FD Frequency difference: The difference between two frequencies expressed in percent, i.e.

$$\mathrm{FD} = 100 \cdot \left(\frac{f_{higher}}{f_{lower}} - 1 \right) \quad [\%].$$

FS Rapidly varying fine structure of a time domain signal.

ITD Interaural time difference: ITDs are the dominant cue for localizing sounds containing low frequencies.

MCL Maximum comfortable level: Maximum current for which the participant accepts permanent stimulation.

MD Modulation depth: Amount of modulation that is introduced to a carrier signal. It ranges between 0 (no modulation) and 1 (max. modulation).

SIPI Short interpulse interval: Describes a special pair of two pulses in a CI pulse train. The term “short“ is not finally quantized in this context.

THR Threshold: Lowest current at which a signal is detected by the participant.

1 Introduction

Pitch is essential for identifying melody, a series of tones, and harmony, a mixture of tones. Further, in speech pitch helps to discriminate questions from statements (prosody) and in tonal languages such as Mandarin Chinese, it even partly defines the meaning of words. In challenging acoustic environments, the so called "cocktail parties" (Cherry 1953), pitch information improves sound source segregation. Pitch is the perceptual equivalent to the periodicity or the repetition rate of a sound (Oxenham 2012).

Cochlear implants (CIs) can at least partially restore auditory perception in the profoundly deaf using predominantly temporal cues. Designed to restore the ability to understand speech, since the 1980s CIs managed to achieve that goal bit by bit, at least for quiet situations (Zeng et al. 2008). Yet, there remain numerous challenges, e.g. tasks involving information on pitch, tonal language processing and also speech perception in noisy environments (Zeng et al. 2005).

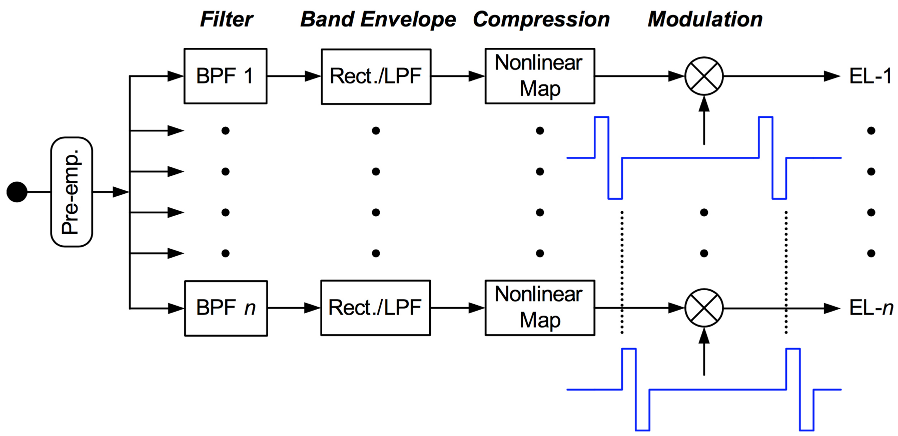


Figure 1.1 – Signal processing in current CIs (Wilson and Dorman 2008).

Most CI processing strategies aim at conveying speech information stored in the temporal envelope (ENV) of the transmitted signals. First, the signals are split up into channels, each allocated to an electrode of the implant (see figure 1.1). Then, in every channel the ENV is extracted, imposed on a periodic pulse train carrier and passed to the corresponding electrode. Thus, the temporal fine structure (FS) of the input signal is discarded.

Rate pitch[1] can be extracted from both FS and ENV, but FS pitch has been shown to be more salient than ENV pitch in normal hearing (Smith et al. 2002). Place pitch[2] information is basically limited to the locations of the electrodes and thus very rough. Hence, CI listeners lack the ability to perceive salient pitch cues and solving this problem remains of major interest in current research.

ITD, Jitter and SIPIs. Laback (2012) summed up recent findings on improved sensitivity to interaural time differences (ITDs) with jitter. Laback and Majdak (2008) found considerable improvements in ITD sensitivity of CI listeners at high pulse rates by applying binaurally coherent jitter to the periodic

1. The term "rate pitch" is used for pitch information based on the temporal signal properties.
2. "Place pitch" refers to pitch based on the tonotopic place of representation in the cochlea.

pulse trains of the CIs as displayed in figure 1.2. Importantly, such high pulse rates are required for robust speech understanding (Loizou et al. 2000, Arora et al. 2009).

The jitter randomly varies the inter-pulse interval (IPI) within a certain range around the IPI of the periodic pulse train. There has been some debate whether the long IPIs or the short IPIs (or both in combination) are the reason for the improved sensitivity. Goupell et al. (2009) tested acoustic pulse trains in normal hearing listeners and reported that the long IPI alone cannot account for the improvements due to jitter. Finally, Hancock et al. (2012) investigated the neural basis of the jitter effect in physiological studies measuring the firing in inferior colliculus neurons of 15 bilaterally cochlear implanted cats. They confirmed the psychophysical results (Laback and Majdak 2008) and further found that particularly pulses with short IPIs, so called SIPIs, are the basis for enhanced ITD sensitivity. Using a threshold model relying on the SIPIs, they were able to accurately predict the neural firing pattern to a 1s-jittered pulse train. Importantly, firing was only restored at the sparse preferred times where the jitter randomly generated the SIPIs.

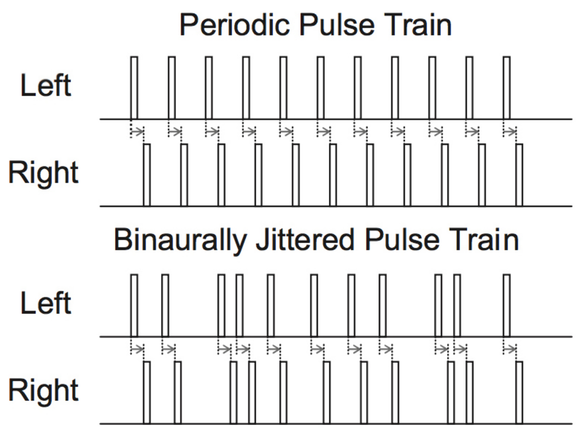


Figure 1.2 – Periodic and binaurally jittered pulse trains containing the same ITD (Laback and Majdak 2008).

It was then hypothesized (Hancock et al. 2012, Laback 2012) that it might be sufficient to modify high-rate pulse trains in a deterministic way by simply inserting extra pulses with SIPIs at certain time instances instead of jittering the whole pulse train. This might improve ITD sensitivity without having to cope with the potential disadvantages of jitter such as lowered speech intelligibility. This hypothesis is currently under investigation at the ARI.

As both ITDs and rate pitch are encoded in the temporal domain, a deterministic way of modifying the high-rate CI pulse trains also opens up new possibilities to transmit rate pitch information. Whereas jitter randomizes the IPIs and thus distorts the carrier information of common CI signals, the SIPI approach as basically shown in figure 2.1 might be useful to encode low-frequency pitch information, such as the F0s of male and female speakers in voiced speech at high carrier pulse rates.

Report Structure. This project report describes the development of an experimental software that allows to test the hypothesis that the introduction of SIPI pulses to common CI signals improves rate pitch perception while maintaining speech intelligibility. It furthermore accounts for potentially confounding influences such as loudness differences between different stimulus conditions.

Additionally, a pilot study is conducted and evaluated to confirm the software's accurate functionality. The data collected should also act as a basis to decide whether an extended study on the effects of SIPI pulses on CI rate pitch perception might be fruitful. Yet, the perceptual evaluation is not part of this project.

2 Pitch SIPI Experiments: Requirements, Design, and Realization

In essence, this project's goal is to design experiments and implement test software for a pilot study that allows to investigate the principal influence of SIPI pulse insertion on rate pitch perception with "pseudo-syllabic" signals in unilateral electric hearing (see also chapter 1). Based on previous findings, the numerous parameters of the experiments will be narrowed down and the setup will hence be justified in general.
The main aspects of the experiment design are the generation of the stimuli depending on several parameters that are either fixed, randomized or varied in the tests, the tasks that the participants have to conduct and the test software that is used.

2.1 Stimuli

The findings of the ITD studies (section 1) leading to the idea of using SIPI pulses to improve the pitch information conveyed by speech-like signals are used to design the stimuli for the Pitch SIPI experiments. Hence, these signals are the actual innovation and form the core of the experiment design.

2.1.1 Signal Terminology

We defined abbreviations for the different signal types which are listed in table 2.1. The parameters are explained below.

signal components	SIPI factor (-)	SIPI phase	abbreviation
AM	— (1) [3]	—	no SIPI
AM+SIPI	1	onset ($\approx 67.5°$)	Full Rate-Onset / FR-O SIPI
		peak ($90°$)	Full Rate-Peak / FR-P SIPI
AM+SIPI	2	onset ($\approx 67.5°$)	Half Rate-Onset / HR-O SIPI
		peak ($90°$)	Half Rate-Peak / HR-P SIPI

Table 2.1 – Signal terminology for the stimuli used in the Pitch SIPI experiments.

2.1.2 Amplitude Modulation and SIPIs

Today's CIs typically use periodic pulse trains as carriers for the information transmitted by the implant (fig. 2.1, top left plot). In our study, the pulse rate was fixed at $f_c = 2$ kHz, which is well in line with previous investigations on envelope rate pitch discrimination (e.g. Galvin et al. 2015, Landsberger

3. For software simplicity reasons, also the pure AM signals have a defined SIPI factor of 1 although the parameter doesn't influence the signal generation.

2008, Vandali et al. 2013). Still, it is rather on the top of previously used carrier rates, therefore the several reasons leading to our decision will be explained during this chapter.

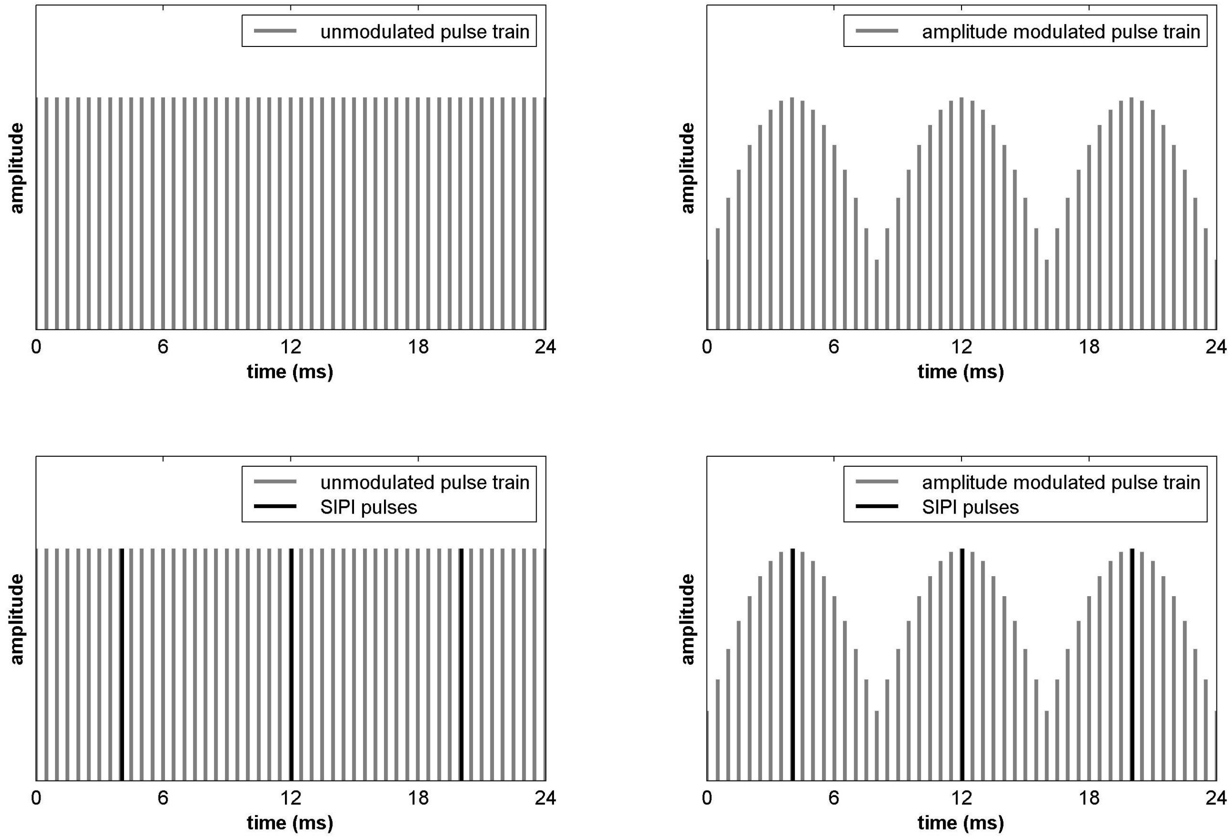


Figure 2.1 – Positive phases of electric pulse trains: *top left:* unmodulated 2-kHz pulse train; *top right:* 2-kHz pulse train, amplitude modulated with a rate of 125 Hz; *bottom left:* unmodulated 2-kHz pulse train with SIPIs inserted with a rate of 125 Hz; *bottom right:* 2-kHz pulse train, amplitude modulation and SIPIs with a rate of 125 Hz.

The audio information extracted from the signals recorded by the CI processors' microphones is imprinted onto the periodic carriers using amplitude modulation (AM, see fig. 2.1, top right plot; also cf. fig. 1.1). In case these audio signals contain some kind of periodicity, e.g voiced speech, the AM is also periodic. For our study, we synthetically produce signals of this type that only contain one constant periodicity information which is expressed as an absolute valued sinusoidal AM with a constant modulation frequency f_{AM}. These signals are referred to as "pseudo-syllabic". Because the goal of the SIPI approach is to enhance F0-cues of male and female speech, our signals contain average F0 information estimated for male (F0 = 125 Hz) and female speech [4] (F0 = 250 Hz), the so called base-F0s. This leads to the first constraint,

$$f_{AM} = \mathrm{F0}, \tag{2.1}$$

4. Proposal of National Insitute of Health (NIH) project "Bilateral Cochlear Implants: Physiology and Psychophysics", grant #: R01DC005775.

which is valid for all signals used in the pilot study. Particularly, the AM of the signals in our study is basically generated according to the following formula:

$$\mathrm{AM} = 1 + \mathrm{MD} \cdot (2 \cdot |\sin[\pi \cdot \mathrm{F0} \cdot t - \phi_0]| - 1), \quad \phi_0 \ldots \text{starting phase.} \tag{2.2}$$

Because the full wave rectification doubles f_{AM}, the frequency of the sine is halved (π instead of 2π). Similar formulas are also used in current ITD experiments at the ARI as well as by Hu et al. (2017) for instance.

In contrast to previous studies that often used only high modulation depths (MD) around 100 % of the participants' DR (comparable to an MD = 1 in our experiments), we use a range of lower MDs. Simulations of CI signals that are produced by typical speech signals of both Med-El and Cochlear Ltd. CIs (Srinivasan et al. 2017) showed that the MD of such signals is way below 1, peaking around 0.25. Thus, we will use two MDs in the pilot study (cf. section 3): to account for the everyday situation, one MD will be at 0.3, the other MD is chosen to be at 0.7 to be able to estimate performance changes for different MDs.

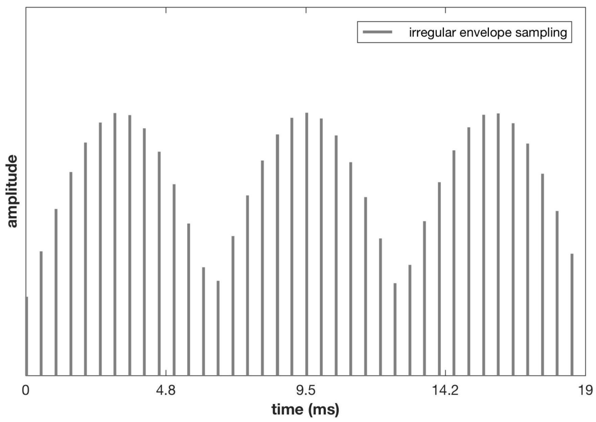


Figure 2.2 – Irregular envelope sampling: if $i = \frac{T_{AM}}{T_c} = \frac{f_c}{f_{AM}} \notin \mathbb{N}$.

Whereas the AM signals act as our reference signals that measure the participants' pitch discrimination sensitivity in present-day situations, our new stimulation strategy modifies these reference signals. Based on the results of both behavioural and neurophysiological studies investigating the jitter effect in electrical pulse trains showing that SIPIs are responsible for improved sensitivity (see chapter 1), we modify the pulse trains by inserting extra SIPI pulses into the static, high rate carriers. This is done in a deterministic way, meaning with a fixed distance to the preceding carrier pulse and a constant rate f_{SIPI} (see fig. 2.1, bottom left plot). The distance between the pulses with the short intervals will be called "SIPI width" and expressed in percent of the carrier pulse interval. As recent investigations have shown (Srinivasan et al. 2017), ITD sensitivity increases with decreasing SIPI width.

The rate of the extra pulses is equal to the F0s corresponding to the AM signal, which will be referred to as Full Rate SIPI insertion:

$$f_{SIPI} = f_{AM} = \mathrm{F0}. \tag{2.3}$$

Additionally, the position of the SIPI insertion is locked within the AM period (see fig. 2.1, bottom right plot). The underlying parameter is called "SIPI phase" and will be varied in the pilot test (see below). This phase criterion in combination with the SIPI width restricts the testable F0s, because they quantize our frequency space quite roughly. We can only test F0s whose corresponding periods are an integer multiple i of the carrier pulse period.

Figure 2.2 illustrates what happens if we choose arbitrary modulation frequencies. Different AM phases are quantized which makes a SIPI pulse insertion with a fixed SIPI width, a constant SIPI phase and $f_{SIPI} = f_{AM}$ impossible. In this context, the decision of the carrier pulse rate is also crucial. The carrier rate f_c directly defines the smallest realizable frequency difference (FD), i.e. the most challenging condition, in the setup of the pitch discrimination experiment.

In contrast to normal hearing where a sampling rate only has to be twice the highest signal frequency, in electric hearing the carrier pulse rate should be four to five times higher than the highest AM frequency to avoid distortions in the neural response to the signals (Wilson et al. 1997). This is another reason that lead us to choose a high carrier rate, i.e. to have enough room to test large FDs even at the high base-F0 of 250 Hz.

2.1.3 Composition of the Pseudo-Syllabic Stimuli

Onset and Offset Ramps. To avoid onset or offset enhancement in the auditory system to influence the results, signals have to be faded in and out. For this and other purposes, the threshold (THR), the maximum comfortable level (MCL) and the comfortable level (CL) have to be measured for individual participants.

The ramps fade the signal in from THR to the defined amplitude of the signal before the steady state and fade it back out to THR after it (see fig. 2.3). They have a fixed length of $t_{ramp} = 150$ ms each and never contain SIPI pulses. For the stimuli used in this study, the amplitude (Amp) is defined as the current of the signal without AM, i.e. MD $= 0$.

Steady State. The steady state basically contains the signals described in the previous section with an AM according to equation 2.2, but with three modifications. Because the AM is centered around the level of the unmodulated pulse train, Amp, there would be an amplitude drop after the onset ramp if the first AM period starts at a SIPI phase of 0° (cf. fig. 2.3). NH listeners perceive this discontinuity as a spectral widening (click), but CI listeners only hear a drop in loudness. Still, this is unwanted and could potentially provide extra cues apart from being disturbing. To solve this issue, we start the first AM period at a phase $\geq 30°$ ($\sin(30°) = 0.5$), which we call starting phase phi_0. Particularly, we determine the phases of the AM signal that are quantized by the pulse train and search for the phase that fits our requirement best. Hence, the first AM period is usually incomplete, its length T_{start} depends on the starting phase which in return depends on the AM frequency as it can be seen in the bottom row of figure 2.3. This fractional AM period is also used in mirrored form as the last AM period before the offset ramp starts. Thus, the stopping phase of the last AM period is 180° minus the starting phase.

Recent studies have investigated the dependency of ITD sensitivity on the place of presenting ITD

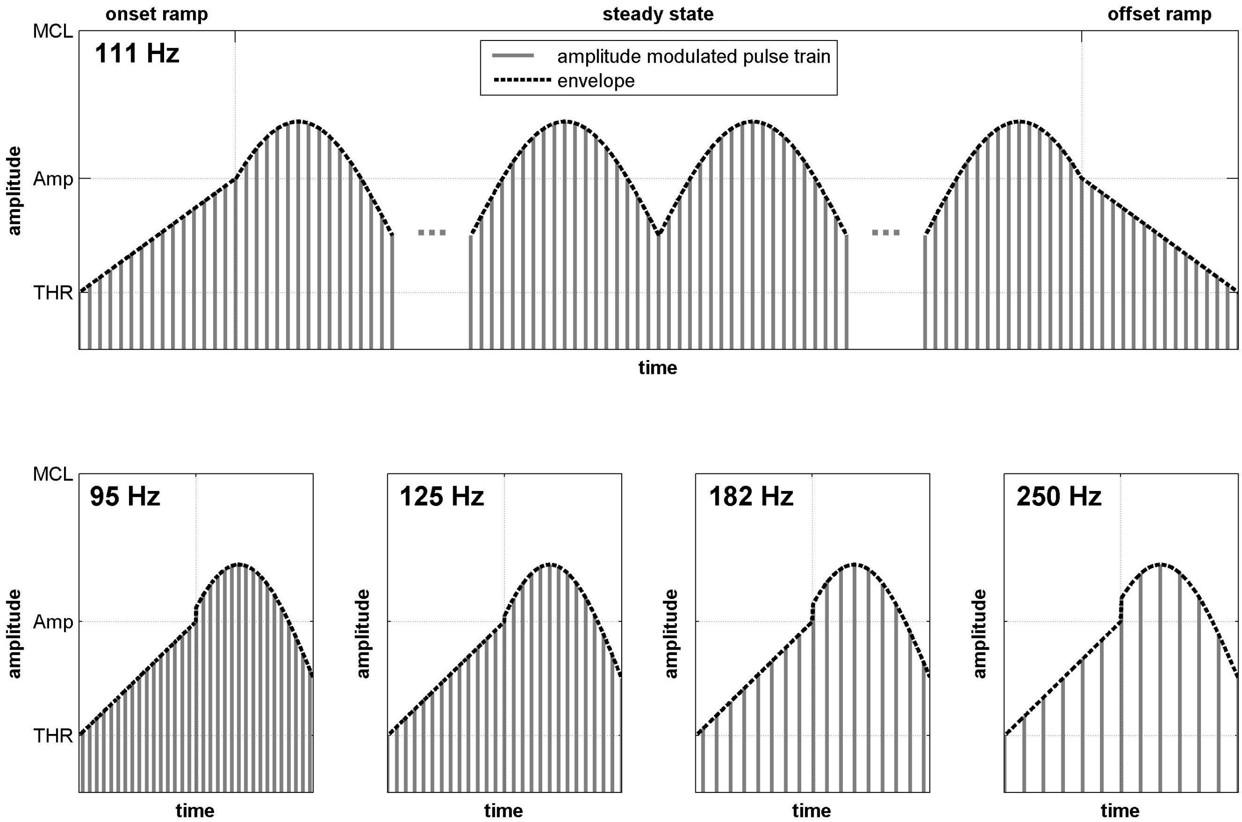


Figure 2.3 – Structure of the stimuli: onset and offset ramp, discontinuities at the beginning of the steady state part with variable length. Depending on the modulation frequency, the starting phase of the first AM period and the stopping phase of the last AM period vary as well as the length of the steady state part.

cues, either at the onset, the peak or the offset of the modulation period (e.g. Hu et al. 2017). While NH listeners seem to be most sensitive to cues at the onset, CI listeners perform better with peak phase cues. Therefore, aiming to quantize the peak of the AM always by a carrier pulse, we had to further modify the stimulus generation. If f_c is an even integer multiple of f_{AM}, the peak is always quantized automatically, but if it is an odd integer multiple, it always lies in the middle of two pulses. As a consequence in this case, we determined the phases which would be represented in the pulse train, shifted them by half the difference between contiguous pulses and then determined the appropriate starting phase using the same criterion as before. By doing so, we now always quantize the peak of the AM, no matter which of our limited set of AM frequencies is chosen.
In the end, we defined that pulses for the onset phase conditions should be inserted at the pulses with quantize the AM phase that is closest to the target onset phase of 67.5°. This is illustrated in figure 2.4. It shows that the actual onset phase is a range around the target phase, table 2.2 contains a selection of possible onset phases for certain AM frequencies used in the pilot study.

Lastly, the overall stimulus length had to be defined. Because the incomplete AM periods vary in

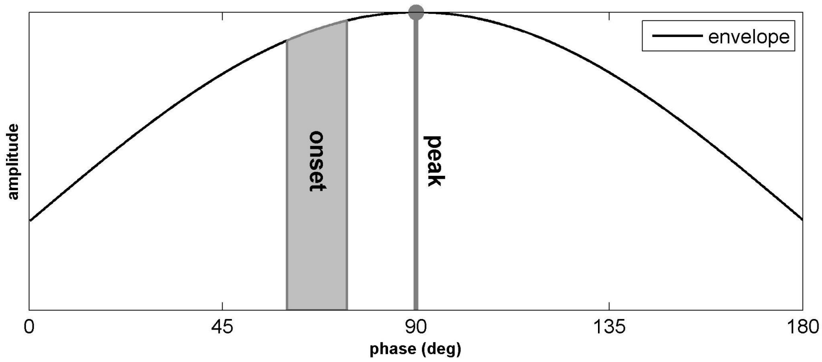


Figure 2.4 – SIPI phase conditions: peak phase (90°) and range of phases between 60° and 74° corresponding to the onset phase condition.

length for different AM frequencies or F0s, respectively, the overall stimulus length couldn't be kept constant for arbitrary AM frequencies as well. Further, if the steady state length would be adjusted for certain F0s based on their F0-specific length of the first and last AM period (remember: the algorithm repeatedly calculates the same starting phase for a certain F0 which then determines the length of the incomplete AM periods), the overall signal length would be linked with the F0, i.e. correlated with it. This could provide additional unwanted length cues in the tasks. To rule this out, we developed the following algorithm: At first, we defined the length of the steady state to be $t_{steady} = 300$ ms. After the shape and thus the length of the two incomplete periods was known, we subtracted their length from the target length. Then, we determined the number of full AM periods (T_0) that fit into the remaining part of the steady state. The remainder of that calculation was then randomly rounded up or down to remove potential length cues. In summary, the steady state has an average length of around 300 ms.

pulse period (ms)	frequency (Hz)	SIPI phase closest to 67.5 (°)
5.5	182	73.6
6.0	167	60.0
6.5	154	62.3
7.0	143	64.3
7.5	133	66.0
8.0	125	67.5
8.5	118	68.8
9.0	111	70.0
9.5	105	71.1
10.0	100	63.0
10.5	95	64.3
11.0	91	65.5
11.5	87	66.5

Table 2.2 – Pulse periods and corresponding onset SIPI phases, $f_c = 2$ kHz.

Overall Stimuli. Combining all considerations explained above, the steady-state signal envelope $f(t)$ (without the ramps) can now be described mathematically, also using formula 2.2, before quantization. To cover different THRs > 0 across participants, the effective MD is adapted:

$$\begin{aligned} f(t) &= \text{Amp} \cdot \text{AM} = \text{Amp} + \text{Amp} \cdot \text{MD} \cdot (2 \cdot |\sin[\pi \cdot \text{F0} \cdot t - \phi_0]| - 1) = \\ &= \text{Amp} + \underbrace{(\text{Amp} - \text{THR})}_{\text{THR} > 0} \cdot \text{MD} \cdot (2 \cdot |\sin[\pi \cdot \text{F0} \cdot t - \phi_0]| - 1), \qquad (2.4) \\ &\quad \text{MD} \in [0,1], \quad t \in [0, t_{max}], \end{aligned}$$

$$t_{max} = 2 \cdot T_{start} + \left(\left\lfloor \frac{t_{steady} - 2 \cdot T_{start}}{T_0} \right\rfloor + \psi \right) \cdot T_0, \quad \psi \sim \mathcal{U}\{0,1\} \qquad (2.5)$$

$$\phi_0 = \underset{k}{\text{argmin}}\ \phi(k) := \left\{ k \in \mathbb{N} \middle| \phi(k) \geq \frac{\pi}{6} \right\}, \qquad (2.6)$$

$$\phi(k) = \begin{cases} k \cdot \frac{\pi}{i}, & i \ldots \text{even} \\ \left(k + \frac{1}{2}\right) \cdot \frac{\pi}{i}, & i \ldots \text{odd} \end{cases}, \quad i = \frac{f_c}{\text{F0}} \in \mathbb{N}. \qquad (2.7)$$

The signal $f(t)$ is then quantized by the carrier pulse train:

$$f_Q(t) = f(t) \cdot \sum_{k=0}^{k_{max}} \underbrace{\delta(t - k/f_c)}_{\text{carrier pulses}} + \text{P}_{SIPI}(t,k), \quad k_{max} = t_{max} \cdot f_c, \quad \delta(t) = \begin{cases} 1, & t = 0 \\ 0, & \text{else} \end{cases} \qquad (2.8)$$

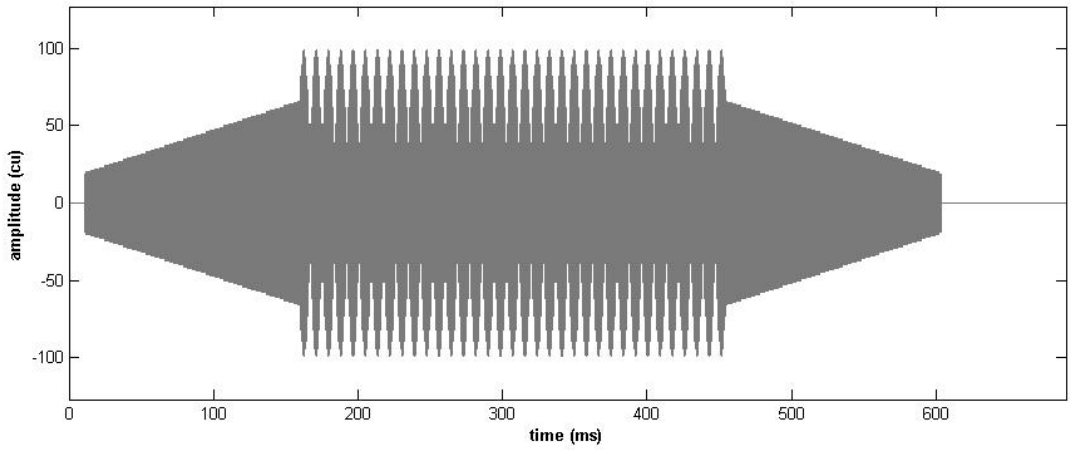
$$\text{P}_{SIPI}(t,k) = \begin{cases} 0, & \text{no SIPI pulses} \\ \delta(t - [k/f_{SIPI} + \text{SIPI width}/(100 \cdot f_c)]), & \text{SIPI pulse insertion} \end{cases} \qquad (2.9)$$

During the pilot study, the signal parameters defining the AM as well as the pulse train are varied as listed in table 2.3.

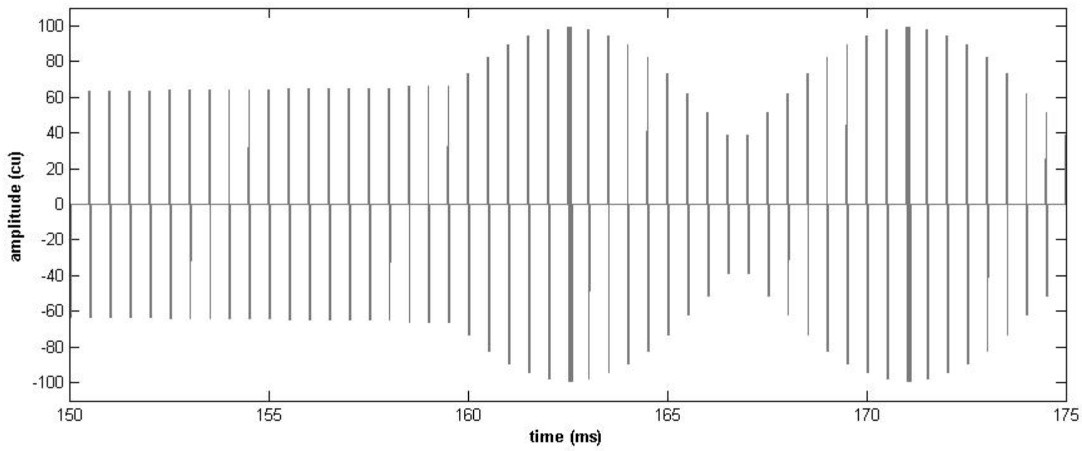
fixed	varied between listeners	varied within listeners
Carrier pulse rate $f_c = \frac{1}{T_c}$	Amp [5]	
Ramp length t_{ramp}	THR	MD
Steady state length t_{steady}		$\text{F0} = f_{AM}$
SIPI width		SIPI rate f_{SIPI}
		SIPI phase

Table 2.3 – Variations of the signal parameters in the pilot study.

5. The amplitudes depend both on the DR of the participants (varied only between them) and the loudness balancing



(a) Entire stimulus.



(b) Zoomed view.

Figure 2.5 – Biphasic pulse train with SIPI pulses: $f_c = 2$ kHz, $f_{AM} = 118$ Hz, SIPI pulses inserted at peak phase.

Unlike other studies where the MD was defined as a proportion of the DR (MCL − THR, e.g. McKay and Henshall 2010, Chatterjee and Oberzut 2011, Vandali et al. 2013, Hu et al. 2017), we defined the MD as the proportion of $2 \cdot (\mathrm{Amp} - \mathrm{THR})$. The reason for this is that we test the participants at their individually reported CL which does not necessarily have to be in the center of the DR. Hence, high MDs would constantly lead to errors. For example, using an MD = 1, there exists only one possible Amp value which would have to be chosen by the listener as CL. Notably, if the participant centers the CL in the DR, our approach is identical with the ones used in the other studies mentioned. In addition it should be mentioned that changing the parameter THR during the study influences the signals crucially and will thus likely lead to errors.

As a final step, after the signal $f_Q(t)$ is generated, the ideal monophasic pulses are converted by the

results (individual).

ExpSuite framework into a biphasic pulse train with a phase duration of 26.7μs (16 temporal units (tu), timebase $1\text{tu} = 1.67\mu$s). The final pseudo-syllabic stimulus as generated by ExpSuite is shown in figure 2.5. This conversion limits the minimum distance between SIPI pulses to $2 \cdot 16$ tu $+ 1 = 33$ tu $\hat{=}$ 55.1 μs which results in a SIPI width of ≈ 11 % for an $f_c = 2$ kHz. Hence, we chose a SIPI width of 12 % for the pilot study.

2.1.4 Half Rate SIPI Condition

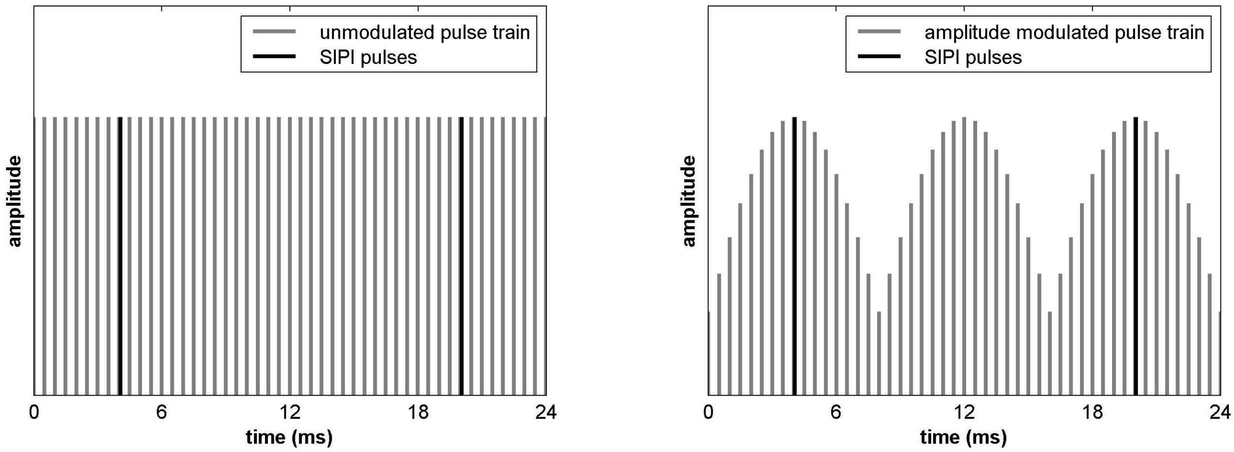


Figure 2.6 – Half Rate SIPI condition: inserted SIPI pulses with half the AM rate.

So far, the SIPI frequency was equal to the F0s. In a second step, As a special condition, the SIPI rate corresonds to half the AM rate (Half Rate SIPI insertion):

$$f_{SIPI} = \frac{f_{AM}}{2}. \tag{2.10}$$

Figure 2.6 illustrates this extra condition. In ExpSuite, which predominantly defines parameters in the time domain (periods), this condition is implemented with the parameter "SIPI factor":

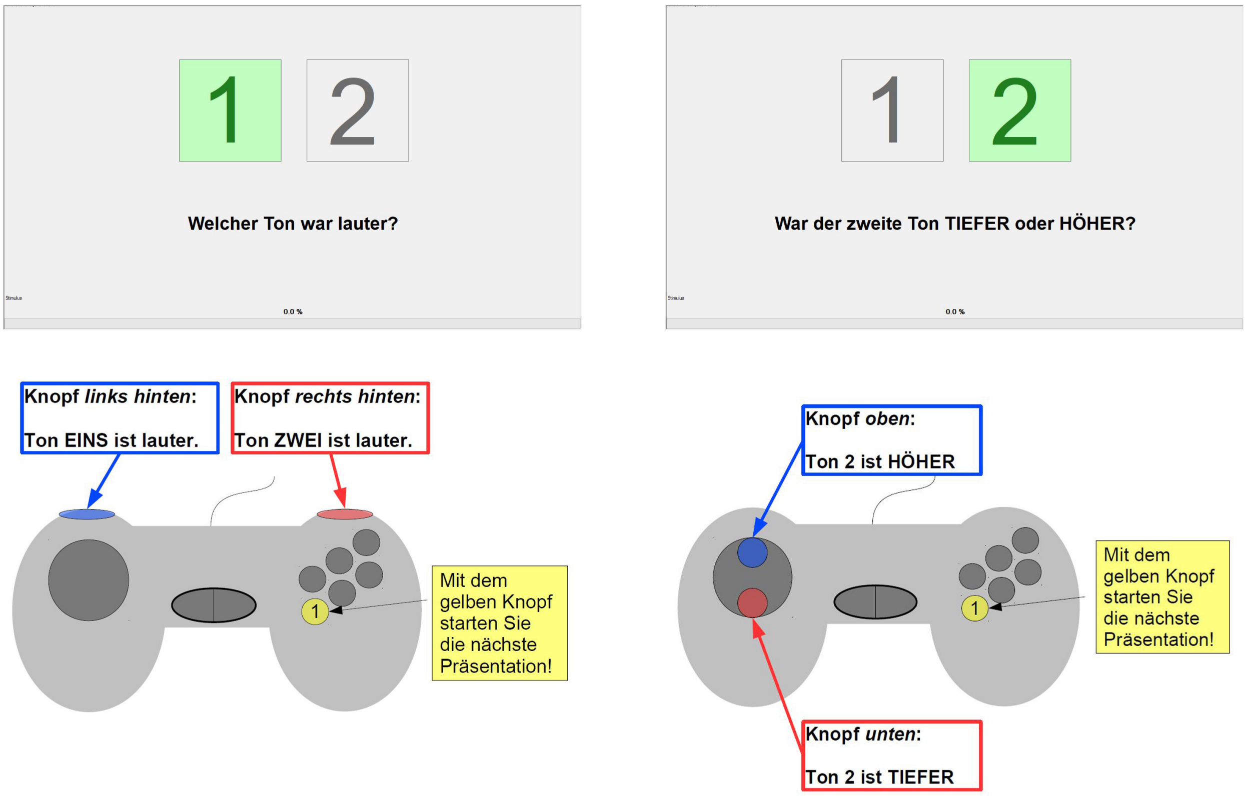
$$T_{SIPI} = \text{SIPI factor} \cdot T_{AM}. \tag{2.11}$$

Obviously, for this condition the SIPI factor equals 2, whereas in the other conditions it is 1.
Using this condition is motivated by the observation in the ITD experiments that the beneficial effect of SIPI pulses declines for SIPI rates f_{SIPI} exceeding 100 Hz. With respect to rate pitch perception, setting the SIPI factor to 2 may downshift the perceived pitch by one octave.

2.2 Task and Procedures

Both for the loudness balancing and the main experiment, a so-called two-interval two-alternative forced choice task (McNicol 1972, 40-45) was used. The listeners' responses were collected using different procedures, based on either the method of limits, the method of constants (or method of constant stimuli), or an adaptive procedure as described by Levitt (1971).

2I-2AFC task. For each judgement, the participants are presented with two signals presented in two intervals (2I) forming a trial. One contains only noise, the other one noise plus the signal. In this context, the signal is referred to as the stimulus property that the participants are asked to detect. They choose one of the two intervals to contain the signal with no possibility to skip responses. Thus, the listeners are confronted with a forced choice with two alternatives (2AFC, McNicol 1972, 40-45).



(a) Loudness balancing task. (b) Pitch discrimination task.

Figure 2.7 – Experiment screens and controller handling for the different experiments in the pilot study.

The presentation of the intervals is visually indicated by highlighting the current interval on the experiment screen (fig. 2.7). Further, the task, i.e. the question the participants have to respond to, and the overall experiment progress are constantly displayed. The participant starts each trial by pushing the yellow button. After the participant's response, the next trial is started automatically. Depending on individual preferences, the length of certain experiment blocks containing a fixed amount of trails can be varied.
Responses are given using a Logitech WingMan® Gamepad. Before the listeners perform an exper-

iment for the first time, they are instructed using a draft of the controller in combination with the instructions which button to push for a particular answer (fig. 2.7). The controller handling is designed to be intuitive to not generate more cognitive load than needed.

2.2.1 Adaptive Loudness Balancing

The adaptive loudness balancing aims at removing potential loudness differences in between the experimental conditions of the pitch discrimination experiment. The measurement procedure is a so called transformed up-down method (Levitt 1971), particularly a 3-down 1-up procedure (Leek 2001).
It is based on the staircase up-down method that adaptively measures the stimulus level x needed to achieve a certain performance level along the psychometric funtion (cf. fig. 2.8). In the case of the simple up-down procedure, this is the performance where the participant has to guess his/her answer. This is normally done by starting at a stimulus level where a very high probability of a positive or correct response, respectively, is estimated. In case of a correct answer, the stimulus level is lowered. This procedure is repeated until the response is negative.
In case one then increases the stimulus level again until a single positive response is obtained, then lowers the level and so forth, this is called a simple up-down or staircase method. The procedure converges at the 50 % point, x_{50}. The x_{50} level can be found by averaging the stimulus levels where a change in level-change direction occurred, the so called turnarounds (Levitt 1971).

Because for 2AFC tasks the chance rate is at 50 %, the x_{50} is not the level of interest in our study. Using a simple up-down method, our loudness balancing procedure would measure nothing but guessing. Hence, we are interested in a stimulus level where the response performance is well above chance, i.e. the level where 79 % percent of the responses are correct, x_{79}. This can be done by transforming the rules of the simple staircase method. Now, the stimulus level is lowered after $n = 3$ correct responses whereas it is increased after every wrong response (3-down 1-up; fig. 2.8 shows a 2-down 1-up method). Thus, the transformed staircase method still converges at the x_{50} level, but this is equal to the x_{79} level of the simple staircase method (Levitt 1971).

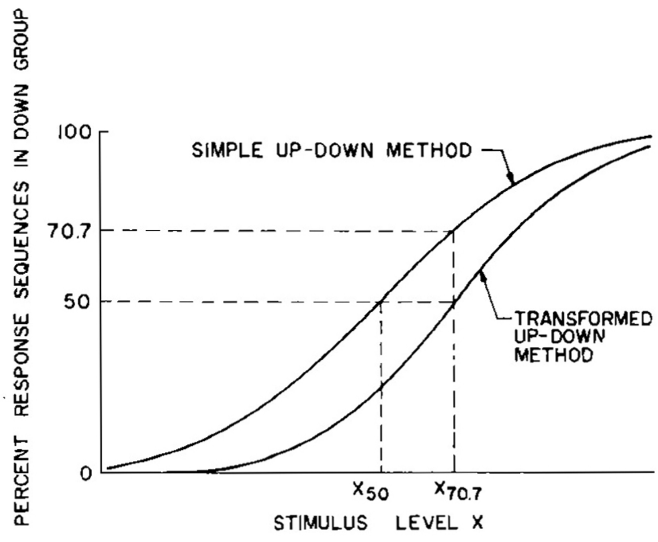
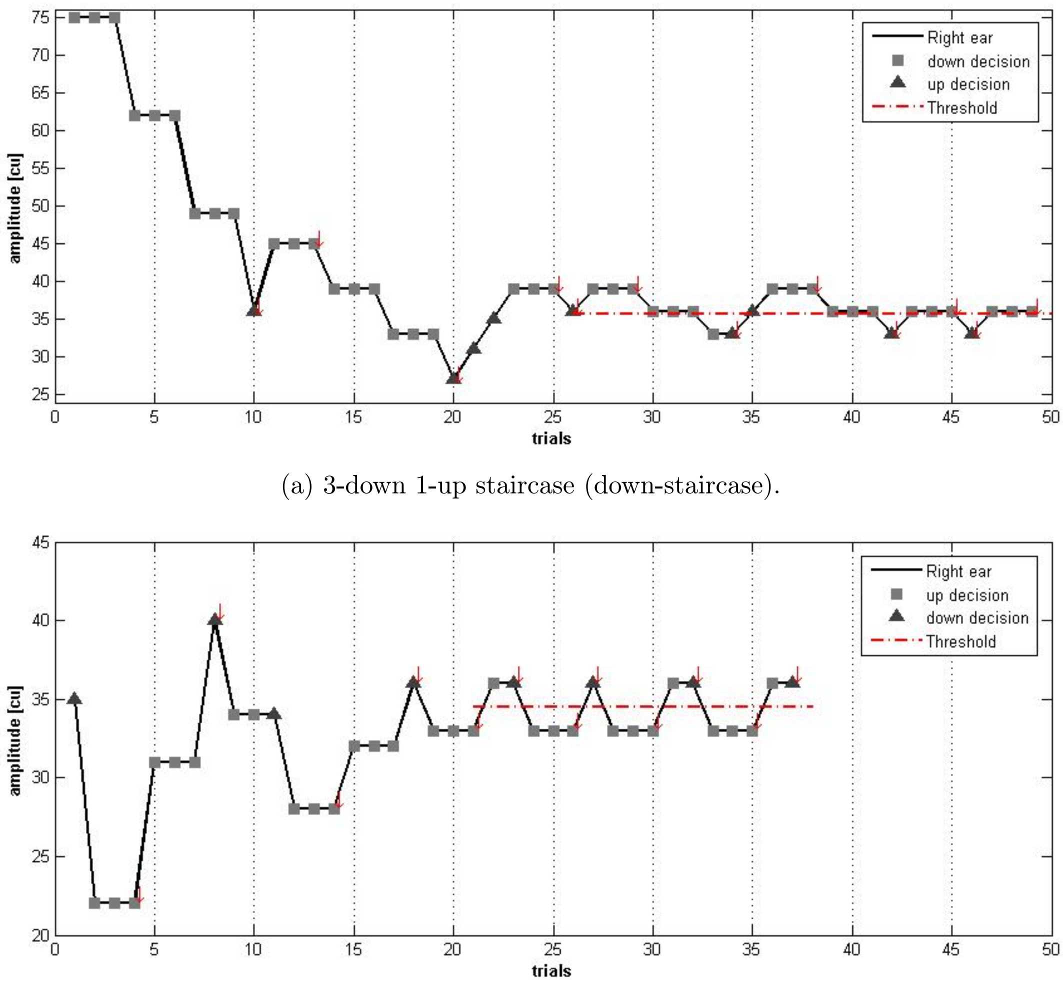


Figure 2.8 – Response curves for a simple up-down and a transformed 2-down 1-up method (Levitt 1971).



(a) 3-down 1-up staircase (down-staircase).

(b) 3-up 1-down staircase (up-staircase).

Figure 2.9 – Adaptive loudness balancing as an example obtained from a listener: MD 0.7/FR-P SIPI to reference (ExpSuite plot).

Implementation in the Pitch SIPI Experiments. Within our experiments, the level of target stimuli is balanced to elicit the same loudness as the fixed level of the reference stimulus which is adjusted to be at the participants' individual CLs. To achieve faster convergence, we modify the size of the steps by which the target amplitudes are changed. Starting at an initial step size μ_0 of 10 % DR, we lower the step size by the factor $1 - d$ after every turnaround (eq. 2.12) until the minimum step size $\mu_{min} = 2$ % DR is reached. From this point, μ stays fixed. The step by which the current amplitude of the target stimulus is changed, μ_{cu} is always computed following equation 2.13.

Moreover, we combine an up- and a down-staircase for every condition. They are described by the decision rule β. In case of the down-staircase ($\beta = 0$), starting from a target amplitude well above the reference, n consecutive responses “target louder“ cause the intensity of the target stimulus to be decreased whereas in case of the up-staircase ($\beta = 1$), starting at a target amplitude far below the

reference, n consecutive responses “reference louder“ cause the intensity of the target stimulus to be increased. This behaviour is denoted in equation 2.14. The target amplitudes cannot be decreased below THR (eq. 2.15) and if they exceed the MCL twice in a row, the item is aborted.
Exemplary, figure 2.9 shows an up- and a down-staircase from the loudness balancing experiments in the pilot study. The termination condition as well as the basis for the result computation can be adjusted in the experiment software and will thus be explained in section 2.3.1.

Adaptive n-down 1-up method:

$$\mu = \max\{\mu_0 \cdot d^{N_T+1}; \mu_{min}\} \quad (\%\ \mathrm{DR}), \tag{2.12}$$

$$\mu_{cu} = \left\lfloor (\mathrm{MCL} - \mathrm{THR}) \cdot \frac{\mu}{100} \right\rceil \quad (\mathrm{cu}), \tag{2.13}$$

$$\mathrm{Amp} = \left\{ \begin{array}{ll} \mathrm{Amp} - (2 \cdot \beta - 1) \cdot \mu_{cu} & \text{Target louder} \\ \mathrm{Amp} + \left\{ \begin{array}{ll} 0, & l < n \\ (2 \cdot \beta - 1) \cdot \mu_{cu}, & l = n \end{array} \right\} & \text{Reference Louder} \end{array} \right\}, \quad \beta \in \{0; 1\} \tag{2.14}$$

$$\mathrm{Amp} = \max\{\mathrm{Amp}; \mathrm{THR}\}. \tag{2.15}$$

<u>Parameters:</u>

μ	...	Step size	μ_{cu}	...	Step in current units
μ_0	...	Initial step size	μ_{min}	...	Minimum step size
d	...	Factor to decrease Step Size	N_T	...	# Turnarounds
β	...	Decision Rule	l	...	# of correct responses in a row
n	...	n-down 1-up method			

2.2.2 Pitch Discrimination

In the pitch discrimination experiments, the participants are presented with two intervals containing stimuli with different F0s. The order of high and low F0 is randomized. The experiments are based on the method of constant stimuli. Because measuring complete psychometric functions is highly time consuming due to the huge number of conditions and because the time with the participants is valuable, we have to to define FDs (cf. fig. 2.8) beforehand (Levitt 1971).
The goal of the experiments is to measure the sensitivity and, in case of influences of SIPI pulses, a change in sensitivity to pitch in the F0s regions corresponding to both base-F0s that we chose (section 2.1). Thus, contrary to the loudness balancing experiments that balanced the amplitudes of several target stimuli to a reference, we are using a different approach here. To ensure that we measure frequency regions that have one of the base-F0s in their center we don't use one of the base-F0s as a

reference and a second, e.g. higher, F0 for the target stimulus as typically done in studies on pitch sensitivity (e.g. Galvin et al. 2015, Ihlefeld et al. 2015, Landsberger 2008).

Instead, we use frequencies that are placed around our base-F0s. In particular, we will use two F0s, $\mathrm{F0}_l < \text{base-F0}$ and $\mathrm{F0}_u > \text{base-F0}$ whose geometric mean F0 corresponds to the base-F0 or the targeted F0 region of interest, respectively:

$$\text{base-F0} = \sqrt{\mathrm{F0}_l \cdot \mathrm{F0}_u}. \tag{2.16}$$

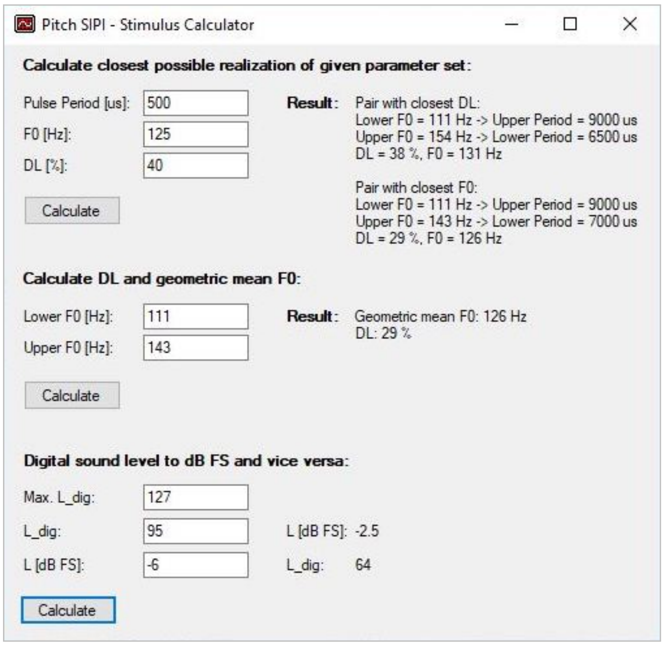


Figure 2.10 – Stimulus Calculator for the PitchSIPI-ExpSuite. Note that FD is called difference limen (DL) in this figure.

The main parameter that changes in the same way for all signal types during the experiments is the MD. Because it is most likely that the participants have different pitch sensitivities, we have to individually adapt the difficulty of the experiment. To be able to observe changes in the pitch sensitivity due to SIPI pulse insertion, we first have to find the FD where the listeners are at the transition from not being able to discriminate the two pitches to perceiving the differences. Ideally, a significant increase in performance, e.g. from close to guessing to good perception, could the be interpreted as an improve in pitch sensitivity.

When the individual FDs are found, the lower and upper frequencies to use in the discrimination experiment can be calculated as follows:

$$\mathrm{F0}_l = \frac{\text{base-F0}}{\sqrt{1+\frac{\mathrm{FD}}{100}}}\ (\mathrm{Hz}) \qquad \text{and} \qquad \mathrm{F0}_u = \text{base-F0} \cdot \sqrt{1+\frac{\mathrm{FD}}{100}}\ (\mathrm{Hz}). \tag{2.17}$$

Due to several constraints that are imposed by the stimulus generation (cf. section 2.1), the frequency range for $\mathrm{F0}_l$ and $\mathrm{F0}_u$ is sampled at integer submultiples of the carrier rate f_c. So it is not possible to realize arbitrary FDs. Moreover, depending on the parameters carrier rate f_c, F0 and FD, there are normally two closest possible realizations: One where the FD is closest to the desired one and the other where the geometric mean frequency of $\mathrm{F0}_l$ and $\mathrm{F0}_u$ is closest to the base-F0. To ease the calculation of possible setups during the experiments, we included a "stimulus calculator" in the pitch discrimination software (see fig. 2.10).

2.3 ExpSuite Applications

The ARI's open source software "ExpSuite" (Mihocic 2014), designed for psychoacoustical tests, is used to implement the Pitch SIPI experiments. As figure 2.11 illustrates, its major advantage is that multiple applications or experiments, respectively, can be realized using the same framework.

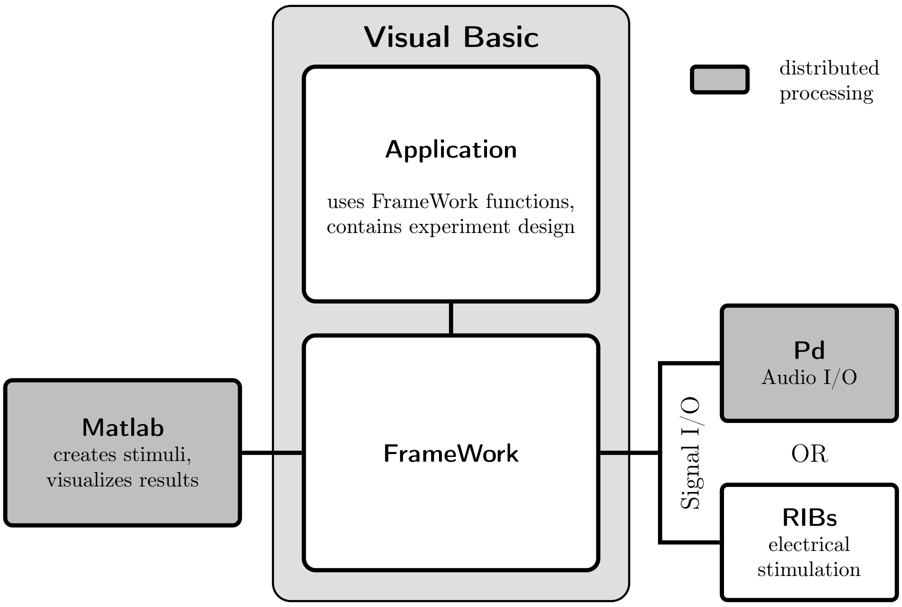


Figure 2.11 – Components of the ExpSuite software for the Pitch SIPI experiments, modified from Mihocic (2014).

Framework and applications are programmed in Visual Basic .NET. Further, the framework also integrates Matlab and Pure Data and provide easy communication functions so that no extra effort by the application developer is needed to gain access to the computational power of these external programs. ExpSuite supports both acoustic stimulation for normal hearing listeners via headphones and electric stimulation for CI listeners using specially developed interfaces called RIBs. Particularly, the Research Interface Box 2 (RIB2, Department of Ion Physics and Applied Physics at the University of Innsbruck, Innsbruck, Austria) is used for the experiments. It allows to bypass the CI processors and directly send stimulation sequences to the implant. The RIB2 supports both unilateral and bilaterally synchronized stimulation.
The software features are pre-distributed between framework and application as listed in figure 2.12.

Framework. The framework defines the software's GUI. This is thus identical for all applications. Nevertheless, the functions of some of the GUI's buttons and Menus can be modified by the application developer. As already mentioned, it communicates with external programs and provides special classes with functions easing the use of the external software.
Apart from that, the framework defines the temporal structure of the stimulus presentation, e.g. pre-stimulus offset, inter-stimulus break and post-stimulus offset. Furthermore, the experiment screen

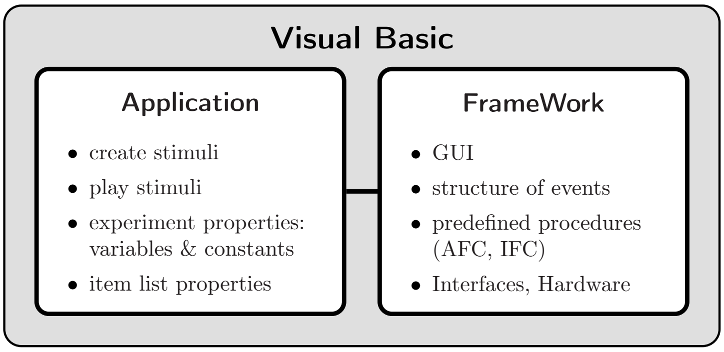


Figure 2.12 – Task distribution in ExpSuite, modified from Mihocic (2014).

and some visual options (e.g. if feedback is provided or not) are implemented for several procedures such as alternative forced choice (AFC). Specific setups can be stored as settings, item lists are saved separately as .csv-files and can be loaded into the application.
In case of electrical stimulation, the software requires so called fitting files that contain the basic sensitivity parameters THR, MCL and CL of the participants for each ear and each electrode.

Applications. The experiment specific features are programmed individually in the applications. Each of them may allow for various experiments. Connecting to the framework features, the developer chooses the procedure for the experiment out of the list of the predefined ones and defines the properties of stimuli used in the experiment. Usually, they are generated using Matlab scripts (see section 2.1). Hence, to specify the signal properties, the developer defines the variables and constants of the experiment(s). Both can be accessed at any time within the software. The variables are varied during the experiment in the item list which is also set up by the developer. Further, though the GUI is fixed, the developer also codes the functions of several buttons that e.g. create item lists, stimulated selected item, start the experiments or evaluate the results based on the participant's responses.
Because the procedures differ substantially between loudness balancing and pitch discrimination (for details see section 2.2), we implemented two separate applications.

2.3.1 ALBforPitchSIPI-ExpSuite

As already noted, the loudness balancing is intended for adjusting the level of multiple experiment conditions so that they are equally loud as the reference signal. Thus, most signal variables exist twice, referring to either the "reference" or the "target" (cf. fig. 2.13). In this ExpSuite, only one experiment type is implemented. The current version number is 1.0.7.
The variables of the experiment can be predefined in the settings (fig. 2.14(a)). There, basic variable properties are listed, e.g. if a variable has to be numeric in general or even an integer. The variable values defined in the settings are used to create item lists using the "Create Liste" button (fig. 2.13). Apart from the experiment variables, the item list also contains columns for the adaptive loudness

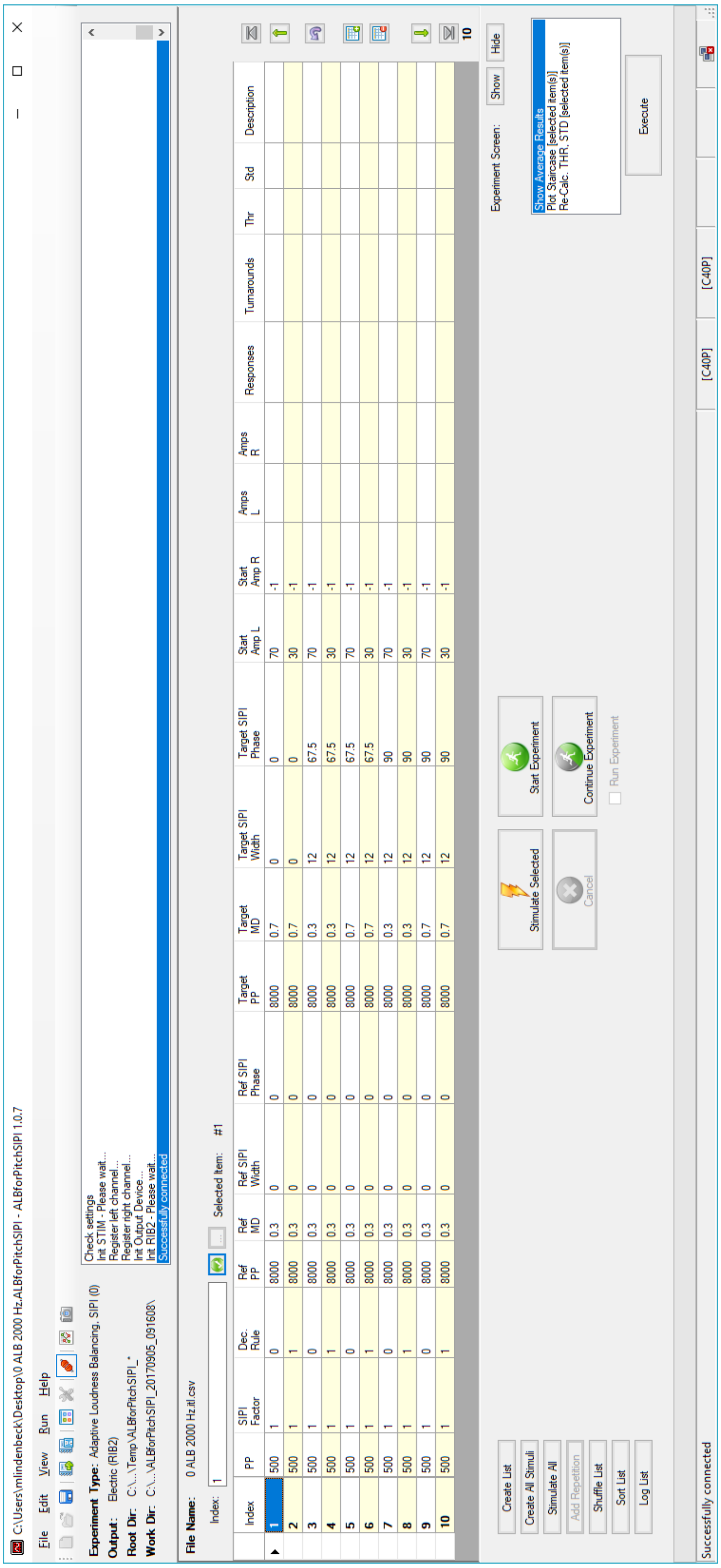
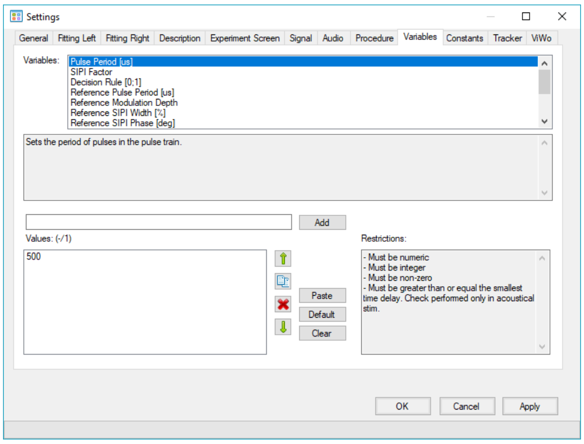
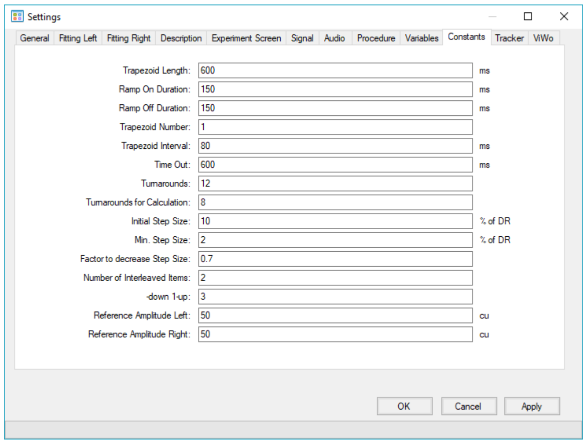


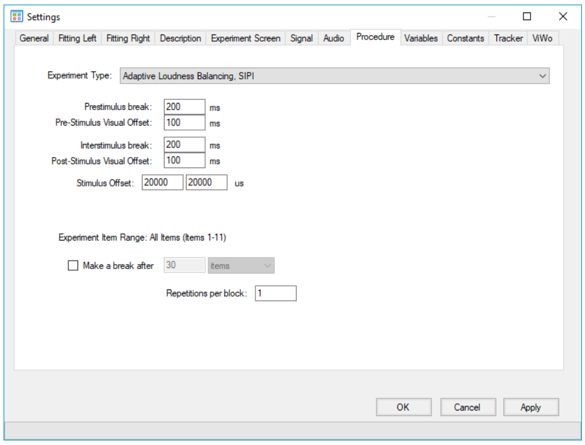
Figure 2.13 – ALBforPitchSIPI-ExpSuite main window.



(a) Variables.



(b) Constants.



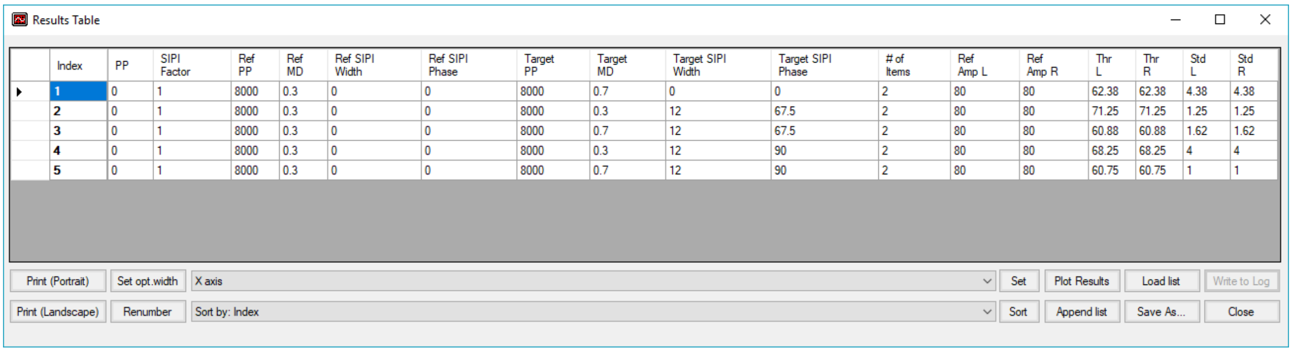
(c) Procedure.

Figure 2.14 – ALBforPitchSIPI-ExpSuite: Application specific settings.

balancing procedure. Because the reference stays constant throughout the experiment, the reference amplitudes are defined in the constants settings (fig. 2.14(b)).
The constants also define the signal length (“Trapezoid Length“), the duration of the ramps, the number of pseudo-syllabic stimuli used in each of the intervals (reference and target; “Trapezoid Number“), the overall length of the signal sequence or interval, respectively (“Trapezoid Interval“) and the parameters of the adaptive method underlying the balancing procedure as summed up in equations 2.12 to 2.15 and the surrounding box.

As also described by equation 2.14, the adaptive method can be used for up- or down-staircases depending on the parameter “Decision Rule“. For our experiments, it is implemented to run every condition twice, an up-staircase and a down-staircase for each condition. Thus, it can be seen in figure 2.13 that every item exists twice but only differs in the value for the decision rule. Further, to hide the algorithm behind the balancing procedure and to reduce order effects, the runs of several items can be

interleaved (constant “Number of Interleaved Items“) randomly. In our case, the two adjacent items describing the up- and down-staircase for one condition are interleaved.
Finally, the length of the balancing procedure and the basis for the result calculation are also defined in the constants. The procedure is aborted after the number of turnarounds specified in the constant “Turnarounds“ is reached and the parameter “Turnarounds for Calculation“ defines how many of the turnarounds, starting from the end, form the basis for the results.
To determine the outcome of the balancing procedure, we implemented three different types of response evaluations: One can either average the matching items, plot the results, or recalculate the results in case some of the procedure parameters, i.e. the number of turnarounds used for the calculation, are changed a posteriori. We average up- and down-staircases as well as repetitions for the same condition. The balanced amplitudes or thresholds (THR), respectively, are calculated by using averaging the amplitudes of the “Turnarounds for Calculation“ (fig. 2.15). Additionally, as an estimate of the participants’ precision, the standard deviation of the averaged results is also computed.



Index	PP	SIPI Factor	Ref PP	Ref MD	Ref SIPI Width	Ref SIPI Phase	Target PP	Target MD	Target SIPI Width	Target SIPI Phase	# of Items	Ref Amp L	Ref Amp R	Thr L	Thr R	Std L	Std R
1	0	1	8000	0.3	0	0	8000	0.7	0	0	2	80	80	62.38	62.38	4.38	4.38
2	0	1	8000	0.3	0	0	8000	0.3	12	67.5	2	80	80	71.25	71.25	1.25	1.25
3	0	1	8000	0.3	0	0	8000	0.7	12	67.5	2	80	80	60.88	60.88	1.62	1.62
4	0	1	8000	0.3	0	0	8000	0.3	12	90	2	80	80	68.25	68.25	4	4
5	0	1	8000	0.3	0	0	8000	0.7	12	90	2	80	80	60.75	60.75	1	1

Figure 2.15 – ALBforPitchSIPI-ExpSuite: Results.

The parameters for the stimulus presentation are defined in the procedure section (fig. 2.14(c)). The parameter “Stimulus Offset“ is a relict from earlier studies and delays the stimulus in the interval by half its value.
Note that “repetition per block“ refers to the number of runs in case of an adaptive procedure.

2.3.2 PitchSIPI-ExpSuite

The software for the pitch discrimination experiments allows to test the listeners ability to discriminate the pitch between two subsequent stimuli of the same type. Hence, the main orientation in the naming of the test variables as well as the organization of the item list is based on the frequency or period, respectively (“upper“ and “lower“). The first stimulus that is created is always the one containing the lower frequency information, but the order of the two signals can be changed using the variable “Order“ which leads either in an upward or a downward movement of pitch across the two intervals.

The ExpSuite includes three experiment types. During the development of the software and also the

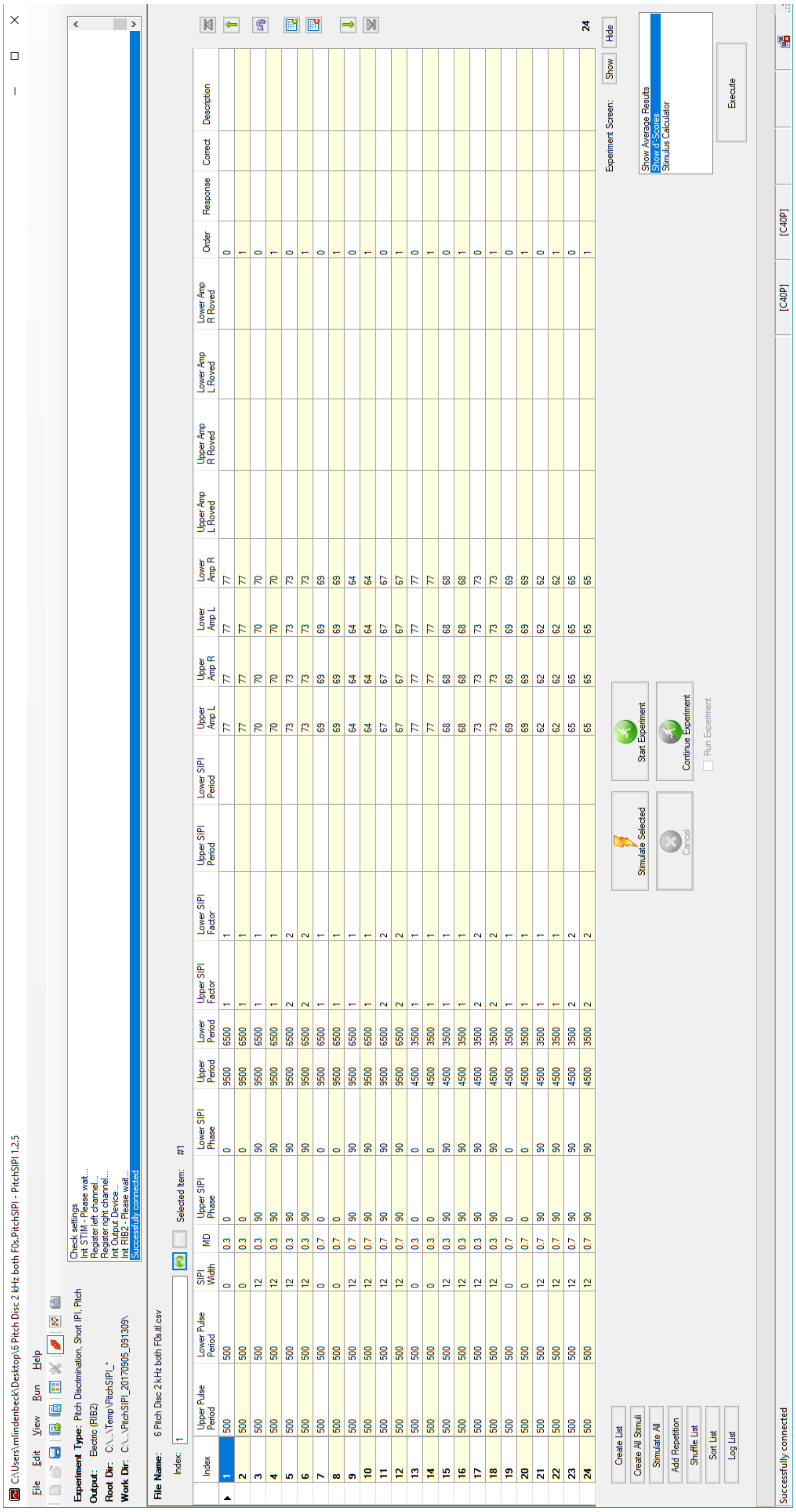
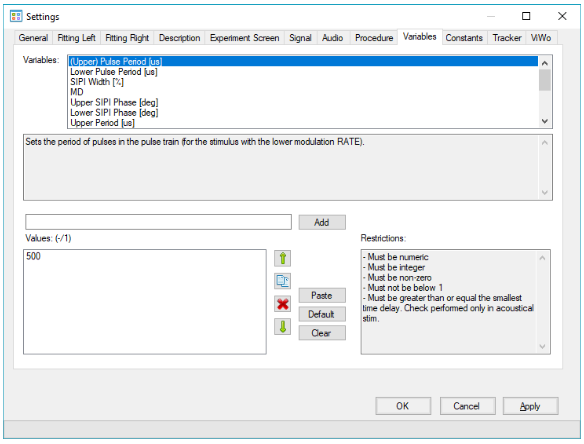
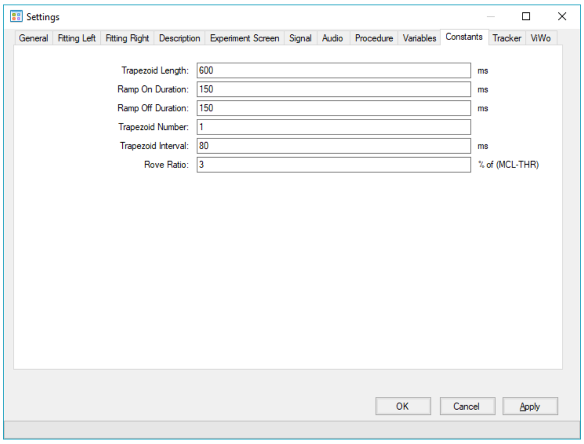


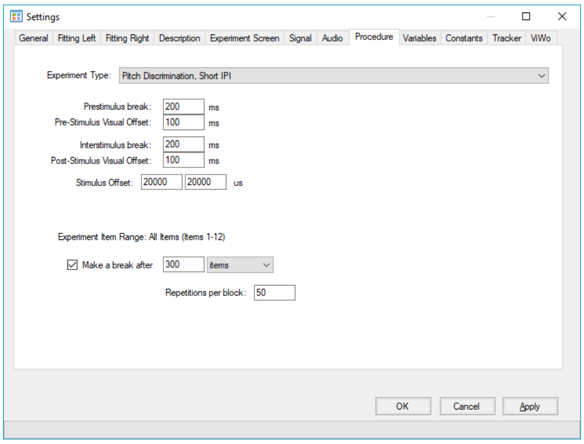
Figure 2.16 – PitchSIPI-ExpSuite main window.



(a) Variables.



(b) Constants.



(c) Procedure.

Figure 2.17 – PitchSIPI-ExpSuite: Application specific settings.

pilot study, more and more variability in the setup of the experiments was needed. Thus, while all types still perform the same experiment and use the same signals, more parameters can be varied between the two intervals. For backward compatibility, the old experiment types are preserved. This, in addition with some bug fixes, leads to a current version of 1.2.5 where figure 2.16 shows the item list for the most extended experiment type.

As in ALBforPitchSIPI-ExpSuite, the values of the (now differing) variables can be predefined in the variable settings (fig. 2.17(a)) and then used for item list creation. Because the experiment uses the method of constant stimuli (section 2.2.2) and thus repeats a fixed set of items a number of times, the constants section of the settings (fig. 2.17(b)) contains much less items than in loudness balancing. They define exactly the same as in the ALBforPitchSIPI-ExpSuite (fig. 2.14(b)), apart from the "Rove Ratio". Although we use the loudness balanced amplitudes in the pitch discrimination experiments, there is the possibility of residual loudness differences. Therefore, random level roving across intervals

of a trial can be included, as specified by the constant “Rove Ratio“ in % DR.
Obviously, also the procedure settings (fig. 2.17(c)) only vary slightly between the two applications. Only the “Repetitions per block“ are much higher, because here the number of repetitions required for the method of constant stimuli are defined. Comparable to the loudness balancing, every conditions is included twice in the item list, only differing in the order of presentation (parameter “Order“). Thus, adding N repetitions leads to a total number of $2 \cdot N$ items of each conditions, with the same number of items for up- and downward movement of the pitch.

In addition to the variables, the item list (again, similar to the ALBforPitchSIPI-ExpSuite) also contains columns for the participants’ response and correct answer flags. To determine the latter, the software analyzes the combination of the variable “Order“ and the response as depicted in table 2.4.

		response	
		0 ↘	1 ↗
order	0 ↘	*correct downward*	wrong
	1 ↗	wrong	*correct upward*

Table 2.4 – Pitch discrimination: Decision logic to determine whether the participant’s response was correct or not. The arrows indicate the movement of the pitch between the two intervals, either as set up by the examiner or as detected by the listener.

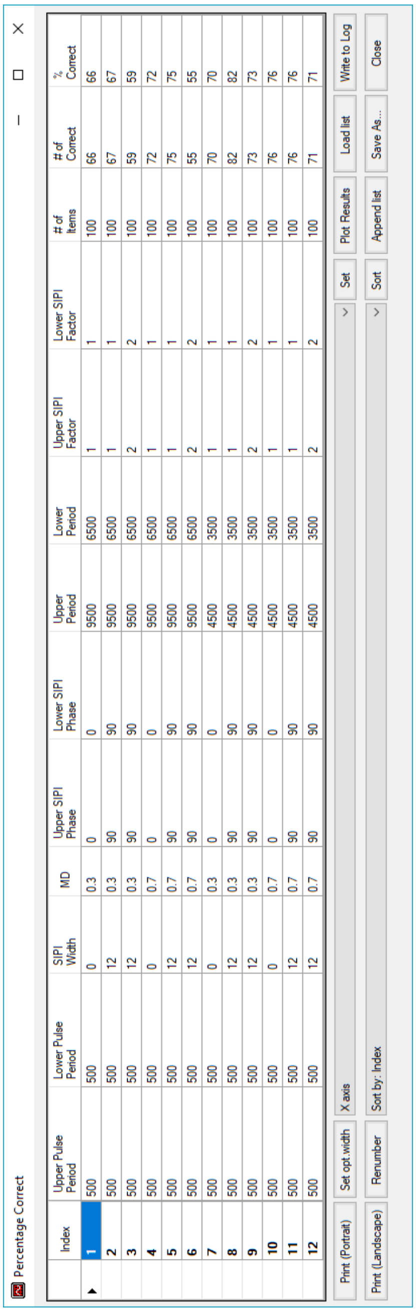
In the item list, the correct answers are not distinguished between upward and downward movement. For the pitch discrimination experiments, we implemented two types of response analysis: Firstly, to determine the participant’s performance without taking into account a possible response bias between up- and downward movement of pitch (“Show Average Results“), the repetitions are pooled and the amount of correct answers leading to a score expressed in terms of % correct responses is counted, as shown in figure 2.18(a). Secondly, for future perceptual evaluations, we also included an evaluation that expresses the results in term of d’-scores following a method proposed by Klein (2001):

$$d' = \frac{z_u + z_d}{\sqrt{2}}, \tag{2.18}$$

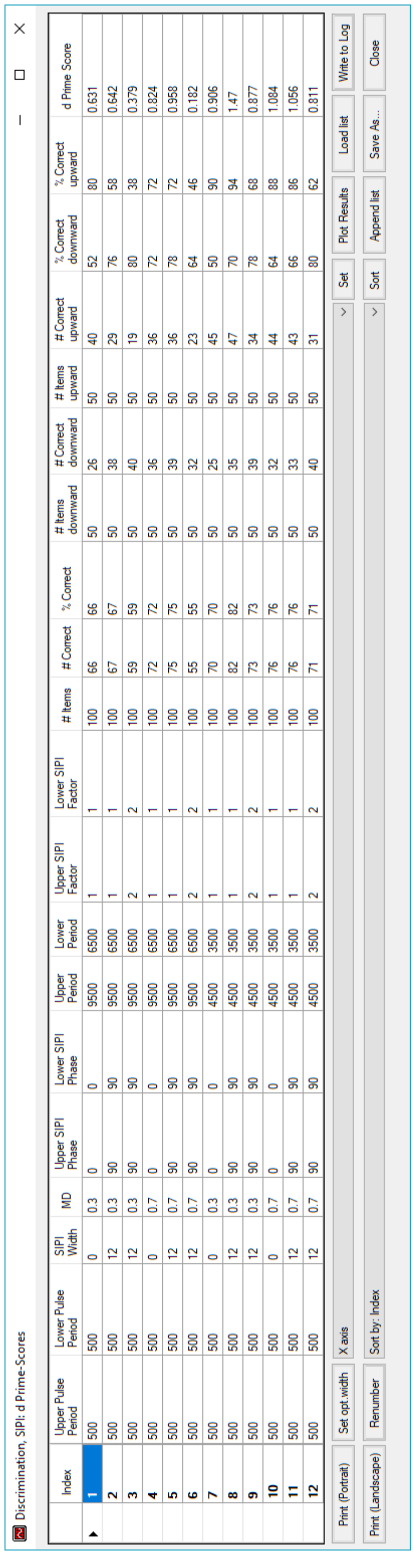
$$z = \Phi^{-1}(P_c) = \sqrt{2} \cdot \mathrm{erfinv}(2 \cdot P_c - 1) \quad \text{(Matlab implementation)}, \tag{2.19}$$

$$P_c = \Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} \mathrm{e}^{-\frac{y^2}{2}}\, dy. \tag{2.20}$$

In contrast to the first evaluation method, which in essence averages the % correct scores P_c for upward and downward movement, the d’ measure averages the z-scores, z_u and z_d of these scores and calculates the d’-values from that (“Show d’-Scores“, fig. 2.18(b)).
Hence, a response bias leads to a decrease in the performance and is thus a more conservative measure. Finally, it should be mentioned here that the previously mentioned stimulus calculator (fig. 2.10) can also be accessed from the results computation section.



(a) Correct responses (%).



(b) d' scores.

Figure 2.18 – PitchSIPI-ExpSuite: Results.

3 Pilot Study

The pilot study is conducted to draw conclusions about the functionality of the experiment software. Apart from the main pitch discrimination experiment itself, the study consists of several subsequent tests that aim at estimating the participants' pitch discrimination sensitivity (training session and pretest) while ruling out loudness cues (loudness balancing).

3.1 Setup

3.1.1 Participants

<table>
<tr><th>ARI Identifier</th><th>CI42</th><th>CI12</th><th>CI18</th><th>CI24</th></tr>
<tr><td>Gender</td><td>female</td><td>female</td><td>male</td><td>female</td></tr>
<tr><td>Age</td><td>73</td><td>51</td><td>60</td><td>55</td></tr>
<tr><td>Implanted ear</td><td rowspan="2">left</td><td>both</td><td rowspan="2">right</td><td>both</td></tr>
<tr><td>Used ear</td><td>right</td><td>left</td></tr>
<tr><td>Electrode #</td><td>3</td><td>2</td><td>8</td><td>8</td></tr>
<tr><td>Implant type [6]</td><td>Pulsar</td><td>C40+</td><td>Pulsar</td><td>C40+</td></tr>
<tr><td>Implant experience (yr)</td><td>17</td><td>17</td><td>13</td><td>13</td></tr>
<tr><td>Deafness duration (yr)</td><td>23</td><td>5</td><td>— [7]</td><td>2</td></tr>
</table>

Table 3.1 – Details about the participants of the pilot study.

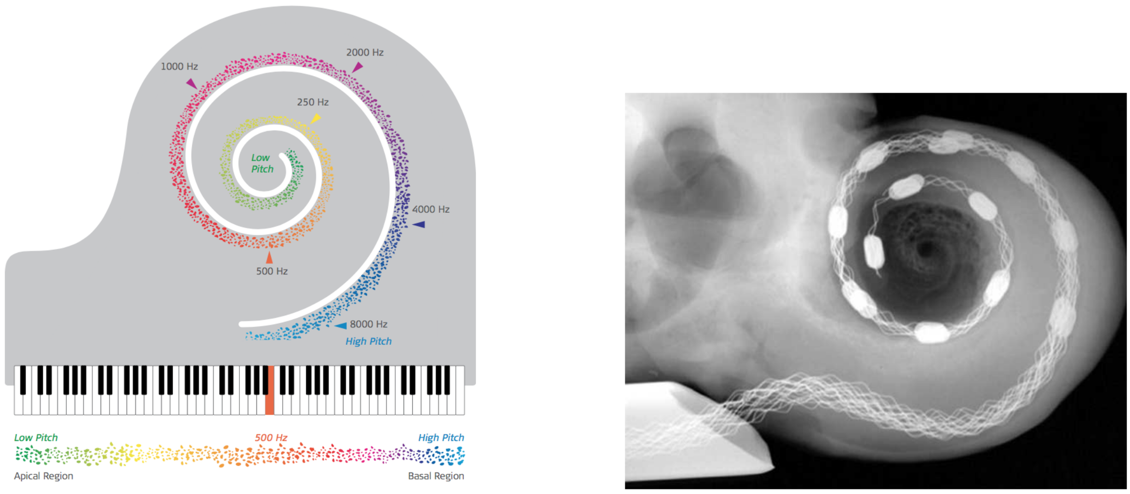
Four post-lingually deafened CI listeners took part in the pilot study, some details of whom are provided in table 3.1. All of them have attended in previous studies and are thus considered as experienced in psychoacoustic experiments. Two participants were unilaterally implanted whereas the other two have bilateral implants. In the latter case, as the study is conducted unilaterally, the listeners were asked to choose their preferred ear.

3.1.2 Electrode Selection

Following the place pitch theory, each CI electrode stimulates neurons in different tonotopic regions of the cochlea. This is demonstrated in figure 3.1 in theory and in practice. The choice of electrode thus determines the place pitch cue. By keeping the electrode fixed troughout the whole experiment, influences of place are kept constant, allowing to investigate the impact of rate pitch.
For the pilot study, we chose two approaches to determine the tested electrode. Following the findings of Stahl et al. (2016), the first two participants were tested at apical electrodes where better rate discrimination performance for low-rate pulse trains ($\leq$ 104 Hz) indicates better temporal sensitivity.

6. Manifactured by Med-EL GmbH, of Innsbruck, Austria
7. The participant reported never to have been completely deaf. Nevertheless, he suffered from severe hearing loss starting approximately at the age of 14 and in the end, his speech recognition scores were below 30 %. He doesn't have residual hearing in the implanted ear, but on the other side (amplified by a powerful hearing aid).

The electrodes for the other two participants were chosen to stimulate the tonotopic place of the carrier frequency f_c used in the experiments ($f_c = 2$ kHz, see section 2.4).



(a) Tonotopic organization in the cochlea (Med-El Corp. 2013).

(b) Location of the implanted electrodes (Dorman and Wilson 2004).

Figure 3.1 – Choice and usage of the implanted electrodes.

According to Baumann and Nobbe (2004), for 12-channel Med-El implants this results in the following decisions: For the first approach, the most apical electrode (#1) should be used and for the second approach, electrode #8 is the electrode of choice. Nevertheless, we evaluated the fitting for the chosen electrodes and in case the dynamic range (DR) was low, we also considered using neighboring, more basal electrodes.

3.1.3 Loudness Balancing

pulse period (ms)	frequency (Hz)	SIPI phase (°)					modulation depth (-)
		no SIPI	Full Rate SIPI		Half Rate SIPI		
			onset	peak	onset	peak	
8	125	reference	67.5	90	67.5	90	0.3
		—	67.5	90	67.5	90	0.7
4	250			90		90	0.3
				90		90	0.7

Table 3.2 – Setup for the adaptive loudness balancing task: reference and target conditions. Only participants CI18 and CI24 conducted the loudness balancing for the 250Hz stimuli.

Vandali et al. (2013) reported loudness changes due to MD variations. Thus, the MD is also a parameter in the loudness balancing experiments. In case of FD, no major influence of the modulation rate of AM stimuli was reported (Chatterjee and Oberzut 2011). Therefore, different F0s were not loudness balanced for signals without SIPI pulses. Besides, for the SIPI conditions, at least separate conditions for the two base-F0s are included to account for extra pulses inserted in the pulse trains.

The setup differed between participants. CI42 and CI12 were only tested at the 125 Hz base-F0. For them, the amplitudes for the 250 Hz base-F0 had to be estimated. CI18 and CI24 were tested at both base-F0s, but for them, the onset phase conditions were omitted because they were excluded from the pitch discrimination experiments. Further, for CI18 and CI24, the whole setup was run twice whereas the first two participants only conducted one run.

3.1.4 Pre-training and Pretest

So far, no data on the ability of our participants to discriminate pitch are collected. Thus, the participants' FDs to be used in the pitch discrimination experiments have to be determined in a pretest. For this, we first calculated all possible FDs ≤ 100 % for a carrier rate of 2 kHz and both base-F0s. Then, we chose certain FDs that we used as our setup for the pretest. The possible FDs as well as the pretest setup are summed up in table 3.3.

pulse periods (ms)	frequencies [geom. mean] (Hz)	FD (%)	used for	
			CI42 CI12	CI18 CI24
8.5/8.0	118/125 [121]	6	P	P
8.5/7.5	118/133 [125]	13		
9.0/7.5	111/133 [122]	20	P	P
9.0/7.0	111/143 [126]	29		
9.5/7.0	105/143 [123]	36		
9.5/6.5	105/154 [127]	46	P	P
10.0/6.5	100/154 [124]	54		
10.5/6.5	95/154 [121]	62		
10.0/6.0	100/167 [129]	67	T	T
10.5/6.0	95/167 [126]	75	P	
11.0/6.0	91/167 [123]	83		
11.5/6.0	87/167 [120]	92		
11.0/5.5	91/182 [129]	100	T	T

(a) 125 Hz base-F0.

pulse periods (ms)	frequencies [geom. mean] (Hz)	FD (%)	used for
			CI18 CI24
4.5/4.0	222/250 [236]	13	P
4.0/3.5	250/286 [267]	14	
4.5/3.5	222/286 [252]	29	P
5.0/3.5	200/286 [239]	43	
4.5/3.0	222/333 [272]	50	
5.5/3.5	182/286 [228]	57	P
5.0/3.0	200/333 [258]	67	
5.5/3.0	182/333 [246]	83	T
6.0/3.0	167/333 [236]	100	

(b) 250 Hz base-F0.

Table 3.3 – Setup for pre-training (T) and pretest (P) as a function of FD.

Further, in figure 3.2 it can be seen how our set of possible FDs and thus frequency pairs systematically fluctuate around the base-F0s and how the frequency pairs spread with increasing FD. At an FD of

100 %, the ranges covered for the two F0s already start to overlap.

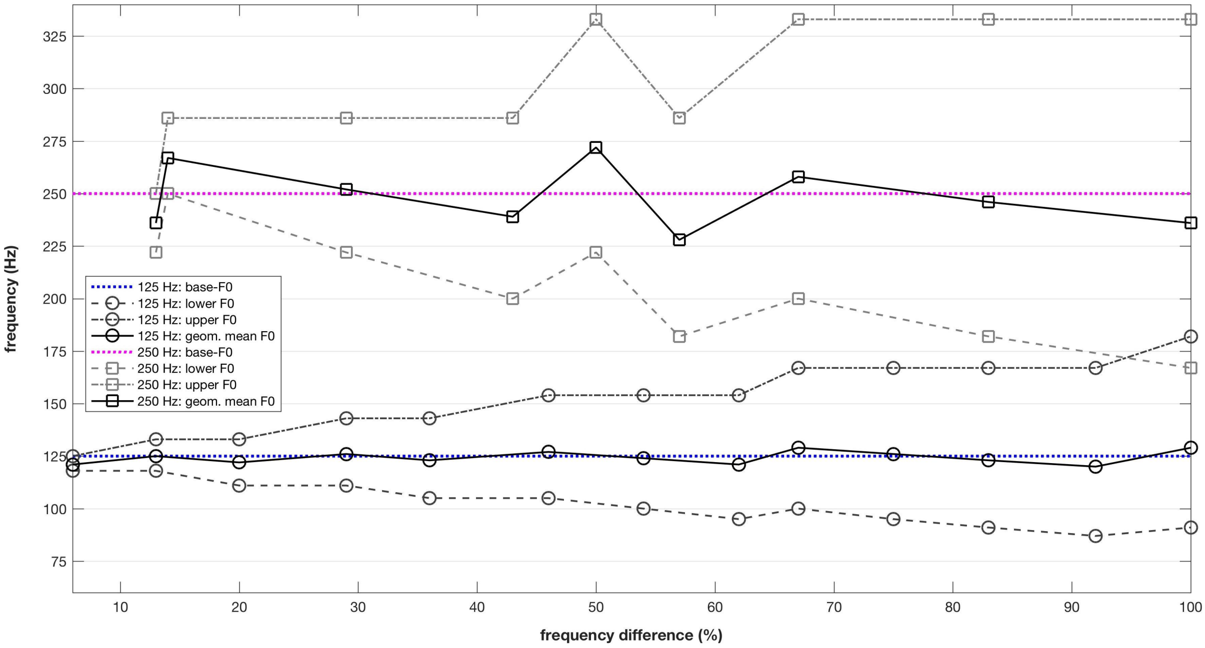


Figure 3.2 – Pitch discrimination: possible frequency pairs and their geometric means for both F0s.

The goal of the pretest was to determine the participants' individual ranges of pitch discrimination performance across all conditions to be expected in the main experiment and thus to avoid floor and ceiling effects. For every FD used in the pretest, we used two conditions in the MD domain (which, again will be the main dimension of the actual pitch discrimination experiments). As the most difficult condition, we selected a low MD (0.3 for CI42 CI12, 0.1 for CI18 and CI24) and pure AM without SIPIs. As for the easiest condition, MD was 0.7 and FR-P SIPIs were used.

As it was the case for the loudness balancing, also the pretest differed between listeners. The first two participants were only tested the 125 Hz base-F0, whereas the other two were examined for both base-F0s. Besides, their setup for the 125 Hz base-F0 was reduced by one FD.
Before the pretesting started, all participants were pre-trained on the stimuli, namely pure AM and FR-P SIPI signals. Using an MD of 0.3 and high FDs, that was assumed to be easy (see tab. 3.3).

3.1.5 Pitch Discrimination

The actual pilot study setup of the main pitch discrimination experiment, denoted in table 3.4, does not differ from the loudness balancing setup apart from the fact that now the frequency information differs in the two intervals. In the loudness balancing, the ideal base-F0s were used for both reference and target. Here, we use one of the frequency pairs in table 3.3 that we decided to take based on the pretests. At this point it should be noted, that any FD from table 3.3 can be chosen and not only

exactly the ones that were used in the actual pretest.
Again, the setup differed slightly between the first two and the last two participants. The onset conditions for the 125 Hz base-F0 were only included in the setup for CI42 and CI12.

pulse periods (ms)	frequencies (Hz) upper / lower	geom. mean	FD (%)	MD (-)	SIPI phase (°) no SIPI	Full Rate SIPI onset	Full Rate SIPI peak	Half Rate SIPI onset	Half Rate SIPI peak
from pretest	from pretest	≈ 125	from pretest	0.3	—	67.5	90	67.5	90
				0.7	—	67.5	90	67.5	90
from pretest	from pretest	≈ 250	from pretest	0.3	—		90		90
				0.7	—		90		90

Table 3.4 – Setup for the pitch discrimination experiments as a function of MD and signal types. The FDs for both F0s were found in the pretest and are now fixed. Test conditions are marked.

3.2 Results

In this section, the results of the loudness balancing and the pitch discrimination tasks are evaluated under the aspect of plausibility in order to verify the implementation of the Pitch SIPI experiments. As it can be seen in table 3.1, due to insufficient DRs at electrode #1 (cf. sec. 3.1.2), the first two participants were tested at electrodes #2 and #3, respectively.

3.2.1 Loudness Balancing

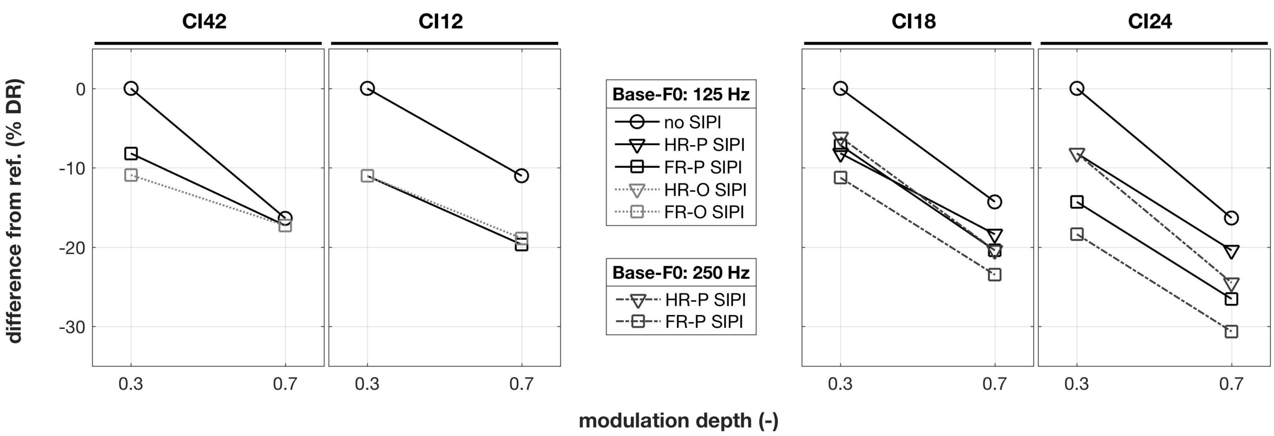


Figure 3.3 – Loudness balancing results as a function of MD for both base-F0s and different signal types, expressed as the difference to the MD 0.3/no-SIPI reference conditions in terms of % DR.

The loudness balancing results are displayed in figure 3.3. All listeners were able to perform the task. Altogether, the results are quite consistent across participants. As the balanced amplitudes (Amp) displayed above determine the level of the pulse train with MD = 0, i.e. the center of the AM, an increase of MD leads to an increase of the peak level (cf. eq. 2.4) and a decrease of the loudness balanced level. Further, the insertion of SIPI pulses increases the signals' energy density and therefore, SIPI signals should be adjusted lower than pure AM signals as it is the case in this pilot study.

Additionally, McKay and Henshall (2010) proposed that the peak amplitudes mainly determine the loudness of the signals. Therefore, to verify the results, the peak amplitudes of the loudness balanced signals were determined and expressed as differences from the reference condition in % DR. The maximum difference between the two MDs, distinguished between different signal types, was 6.7 % DR, the mean difference of all signal types 0.7 % DR and mean squared error (MSE) adds up to 2.1 % DR. This is well in line with McKay and Henshall (2010) and thus seen as plausible.

For participants CI42 and CI12, due to time restrictions not all conditions run in the pitch discrimination experiments (cf. fig. 3.4) were loudness balanced. For the conditions were no loudness data were collected, the amplitudes were extrapolated based on the assumptions on loudness made in section 2.

3.2.2 Pitch Discrimination

Essentially, two dimensions of the test setup are varied during the pilot: Firstly, in the pretest, we search for a challenging FD (thus vary it) while picking out certain modulation depths and signal types. Secondly, during the actual test, the FD is fixed at each base-F0 and we vary the MD and the signal types. Hence it is useful to plot the results both as a function of the FD and the MD.

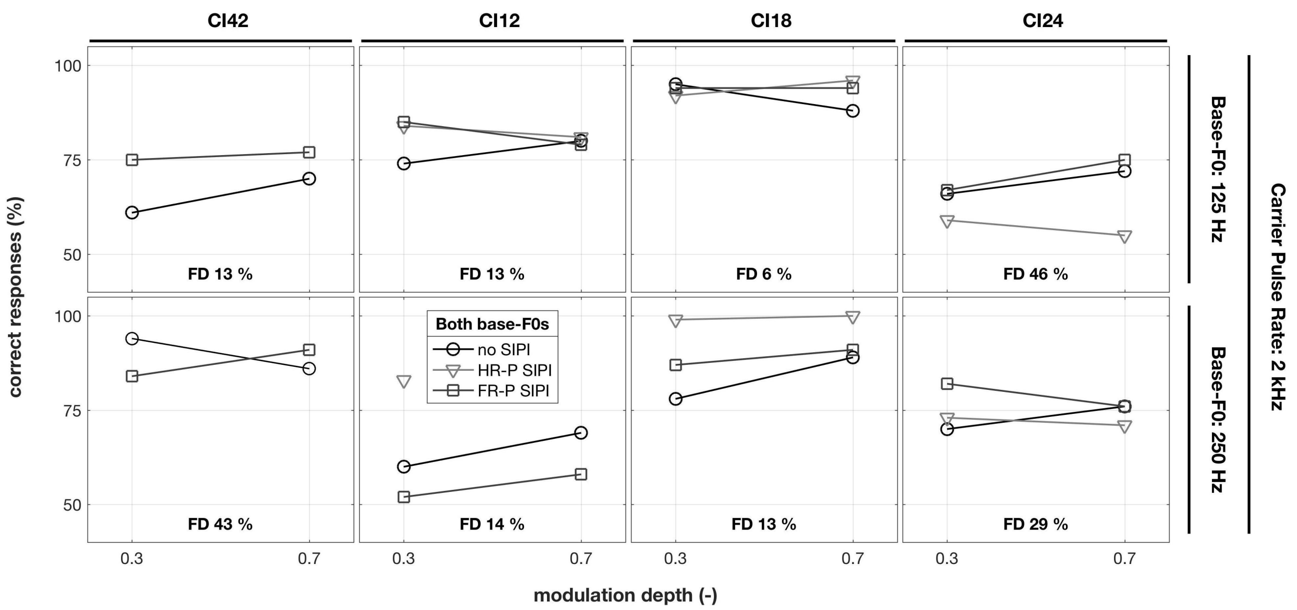


Figure 3.4 – Pitch discrimination results as a function of MD.

Results as a function of MD. Figure 3.4 shows the pitch discrimination results of all four participants and for both F0s. Apart from the results of CI42 (250 Hz base-F0) and CI18 (125 Hz base-F0), the performance for pure AM signals without SIPI pulses is plausible following the hypothesis that an increased MD leads to a better coding of the F0 based on sharper peaks in the signal. Yet, the results for AM+SIPI signals, both HR-P and FR-P, are not as consistent. This might be due to changes in the signal properties caused by the introduction of the extra pulses. Since the evaluation conducted in this report does not include the discussion of perceptional changes, the results of the pilot study are considered reasonable based on the AM signal performances.

Although being included in the loudness balancing task for the first two listeners, the pitch discrimination results do not include conditions with SIPI pulses inserted at the onset. Despite running the pretest to estimate the individual's range of pitch discrimination sensitivity, we didn't choose the "correct", adequately challenging FD at the first try for the first two participants. Hence, we ran the pretest for three different FDs for CI42 and for two FDs for CI12. The results for these additional runs are displayed in figure 3.5. As a consequence, we ran out of time for the last, main runs.

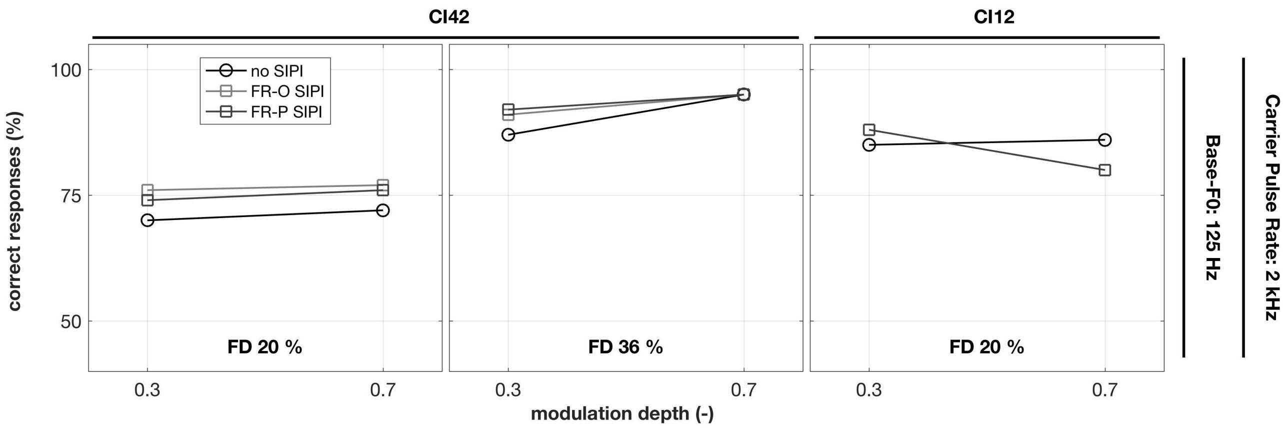


Figure 3.5 – Pitch discrimination results, additional runs as a function of MD.

In these extra runs, the overall picture of performance is similar to that of the main runs. Moreover, the additional runs also included conditions with onset insertion but, because the outcome didn't differ substantially from the peak conditions, they were not included in the following rounds and also for the last two participants.

Results as a function of FD. Having carried out the additional pitch discrimination runs, the results can be evaluated along a second dimension, the FD. An increase in FD is assumed to ease the task and thus increase the performance. Figure 3.6 shows monotonically increasing performances for all signal types, MDs and listeners.
This observation further confirms the basic functionality of the Pitch SIPI experiment software as well as the overall study procedure.

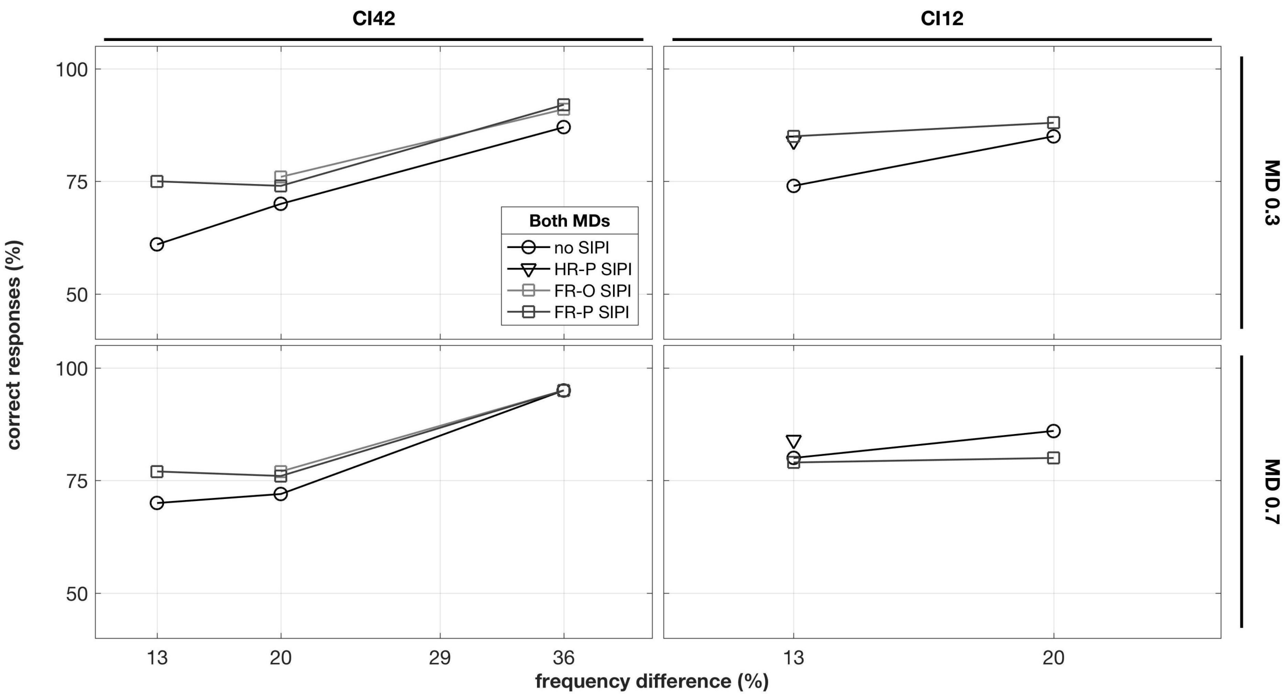


Figure 3.6 – Pitch discrimination results as a function of FD.

4 Discussion and Outlook for Main Study

Software Functionality. The evaluation of the pilot study results in chapter 3 suggests that the ALBforPitchSIPI-ExpSuite as well as the PitchSIPI-ExpSuite indeed implement the study design as described in chapter 2. Apart from individual data, no unexpected overall results were observed and all participants were able to perform both the loudness balancing and the pitch discrimination task. The setup for the pitch discrimination task was sufficient for all participants meaning that it provided the appropriate range of sensitivities for the different conditions.

Study Procedure. The sequence of (i) adaptive loudness balancing, (ii) pre-training, (iii) pretest and (iv) pitch discrimination was efficient and expedient. After the necessary loudness balancing, the pre-training already allowed a first appraisal of the participant's overall sensitivity to pitch and its discrimination. Subsequently, the Pretest provided, at least after some practice for the experimenter, a sufficient estimate of the listener's actual range of sensitivities. This allowed to choose a fixed FD for the pitch discrimination task.

Yet, two aspects seem improvable: First, the pre-training didn't include signals with SIPIs which might lead to an underestimation of performance in the pretest for the SIPI conditions. Second, for both pre-training and pretest, including conditions for the 250 Hz base-F0 seems feasible.

Outlook. For an extended main study following this pilot investigation, some modifications and additions seem natural:

(i) *MD:* As performance changes across MDs for both tasks, more MDs could be included.

(ii) *SIPI phase:* Recently, Hu et al. (2017) showed that CI listeners are most sensitive to ITD cues inserted at the peak of the modulation which also matches the present results. Thus, a main study might focus on peak SIPI insertion.

(iii) *Carrier pulse rate:* The 2 kHz carrier rate used in the pilot is high, but not unusual. Still, including more, maybe lower carrier rates might complement the main study.

(iv) *Overall effects:* In case the individual results will be combined to search for overall effects, the variances of the % correct scores should be homogenized using the rationalized arcsine transform (Studebaker 1985) before running inferential statistics such as ANOVAs.

(v) *Low rate unmodulated (LRU) stimuli:* LRU pitch discrimination performance is well established as summarized e.g. by Vandali et al. (2013). Including these signals might act as a useful benchmark for pitch discrimination sensitivity with pure AM as well as SIPI signals.

(vi) *HR-P SIPI conditions:* The results for these conditions show the biggest variance and are not consistent between participants, i.e. CI12 and CI18 show improvements or no change whereas CI24 shows a performance drop. More detailed investigations appear to appropriate here.

References

Arora, K., Dawson, P., Dowell, R. and Vandali, A. (2009), 'Electrical stimulation rate effects on speech perception in cochlear implants', *Int J Audiol* **48**, 561–567.

Baumann, U. and Nobbe, A. (2004), 'Pitch Ranking with Deeply Inserted Electrode Arrays', *Ear Hearing* **25**, 275–283.

Chatterjee, M. and Oberzut, C. (2011), 'Detection and rate discrimination of amplitude modulation in electrical hearing', *J Acoust Soc Am* **130**(3), 1567–1580.

Cherry, E. C. (1953), 'Some Experiments on the Recognition of Speech, with One and with Two Ears', *J Acoust Soc Am* **25**(5), 975–979.

Dorman, M. F. and Wilson, B. S. (2004), 'The Design and Function of Cochlear Implants', *Am Sci* **92**, 436–445.

Galvin, J. J., Oba, S., Başkent, D. and Fu, Q.-J. (2015), 'Modulation frequency discrimination with single and multiple channels in cochlear implant users', *Hearing Res* **324**, 7–18.

Goupell, M. J., Laback, B. and Majdak, P. (2009), 'Enhancing sensitivity to interaural time differences at high modulation rates by introducing temporal jitter', *J Acoust Soc Am* **126**(5), 2511–2521.

Hancock, K. E., Chung, Y. and Delgutte, B. (2012), 'Neural ITD coding with bilateral cochlear implants: effect of binaurally coherent jitter ', *J Neurophysiol* **108**(3), 714–728.

Hu, H., Ewert, S. D., McAlpine, D. and Dietz, M. (2017), 'Differences in the temporal course of interaural time difference sensitivity between acoustic and electric hearing in amplitude modulated stimuli', *J Acoust Soc Am* **141**(3), 1862–1873.

Ihlefeld, A., Carlyon, R. P., Kan, A., Churchill, T. H. and Litovsky, R. Y. (2015), 'Limitations on Monaural and Binaural Temporal Processing in Bilateral Cochlear Implant Listeners', *J Assoc Res Otolaryngol* **16**(5), 641–652.

Klein, S. A. (2001), 'Measuring, estimating, and understanding the psychometric function: A commentary', *Percept Psychophys* **63**(8), 1421–1455.

Laback, B. (2012), 'Neural basis of improved ITD sensitivity with jitter', *J Neurophysiol* **108**(3), 712–713.

Laback, B. and Majdak, P. (2008), 'Binaural jitter improves interaural time-difference sensitivity of cochlear implantees at high pulse rates', *P Natl Acad Sci USA* **105**(2), 814–817.

Landsberger, D. M. (2008), 'Effects of modulation wave shape on modulation frequency discrimination with electrical hearing', *J Acoust Soc Am* **124**(2), EL21–EL27.

Leek, M. R. (2001), 'Adaptive procedures in psychophysical research', *Percept Psychophys* **63**(8), 1279–1292.

Levitt, H. (1971), 'Transformed Up-Down Methods in Psychoacoustics', *J Acoust Soc Am* **49**(2), 467–477.

Loizou, P. C., Poroy, O. and Dorman, M. (2000), 'The effect of parametric variations of cochlear implant processors on speech understanding', *J Acoust Soc Am* **108**(2), 790–802.

McKay, C. M. and Henshall, K. R. (2010), 'Amplitude Modulation and Loudness in Cochlear Implantees', *J Assoc Res Otolaryngol* **11**(1), 101–111.

McNicol, D. (1972), *A Primer of Signal Detection Theory*, Lawrence Erlbaum Associates, Inc.

Med-El Corp. (2013), 'Understanding Cochlear Implants', medel.com.
URL: *http://s3.medel.com/downloadmanager/downloads/maestro_2013/en-GB/20329.pdf*

Mihocic, M. (2014), ExpSuite: Open Source-Framework zur Durchführung psychoakustischer Experimente und akustischer Messungen, Master's thesis, Fachhochschule Technikum Wien.

Oxenham, A. J. (2012), 'Pitch Perception', *J Neurosci* **32**(39), 13335–13338.

Smith, Z. M., Delgutte, B. and Oxenham, A. J. (2002), 'Chimaeric sounds reveal dichotomies in auditory perception', *Nature* **416**(6876), 87–90.

Srinivasan, S., Laback, B. and Majdak, P. (2017), 'Improving interaural time difference sensitivity using short interpulse intervals with vowel-like stimuli in bilateral cochlear implants', *The Journal of the Acoustical Society of America* **141**, 3973–3974.

Stahl, P., Macherey, O., Meunier, S. and Roman, S. (2016), 'Rate discrimination at low pulse rates in normal-hearing and cochlear implant listeners: Influence of intracochlear stimulation site', *J Acoust Soc Am* **139**(4), 1578–1591.

Studebaker, G. A. (1985), 'A "Rationalized" Arcsine Transform', *J Speech Lang Hear R* **28**, 455–462.

Vandali, A., Sly, D., Cowan, R. and van Hoesel, R. (2013), 'Pitch and loudness matching of unmodulated and modulated stimuli in cochlear implantees', *Hearing Res* **302**, 32–49.

Wilson, B. S. and Dorman, M. F. (2008), 'Cochlear implants: a remarkable past and a brilliant future', *Hearing Res* **242**(0), 3–21.

Wilson, B. S., Finley, C. C., Lawson, D. T. and Zerbi, M. (1997), 'Temporal representations with cochlear implants', *Am J Otol* **18**, 30–34.

Zeng, F.-G., Nie, K., Stickney, G. S., Kong, Y.-Y., Vongphoe, M., Bhargave, A., Wei, C. and Cao, K. (2005), 'Speech recognition with amplitude and frequency modulations', *P Natl Acad Sci USA* **102**(7), 2293–2298.

Zeng, F.-G., Rebscher, S., Harrison, W. V., Sun, X. and Feng, H. (2008), 'Cochlear Implants: System Design, Integration, and Evaluation', *IEEE Rev Biomed Eng* **1**, 115–142.

List of Figures

List of Tables