Master’s Thesis

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

Temporal Pitch in Electric Hearing with Short-Interpulse-Interval Stimulation

066 413 Master Programme Electrical Engineering and Audio Engineering
University of Technology and University of Music and Performing Arts, Graz

Martin Lindenbeck, BSc, 01231382

Advisor:	Doz. Dr. Piotr Majdak
Co-Advisor:	Doz. Dr. Bernhard Laback
Co-Advisor:	Prof. Dr. Robert Höldrich

May 28, 2018

Institute for Electronic Music and Acoustics

Acoustics Research Institute

Abstract

Cochlear implants (CIs) partially restore the auditory perception of people with profound hearing loss by electrically stimulating the auditory nerve using an array of implanted electrodes. Temporal information like pitch and interaural time differences (ITDs) is only imperfectly transmitted via CIs because most of the signal encoding strategies discard the fine structure of the audio signal. Additionally, both pitch and ITD sensitivity are hampered by a rate limitation at high carrier pulse rates needed for speech intelligibility. Recently, a new encoding strategy based on inserting extra pulses with short-interpulse-intervals (SIPIs) to periodic carrier pulse trains has been shown to increase CI-ITD sensitivity. Within this master's thesis, a psychoacoustical study with six CI listeners was conducted to investigate the effect of two SIPI pulse insertion approaches on temporal pitch discrimination sensitivity. Pseudo-syllable signals were used. These signals mimic voiced speech segments by coding the fundamental frequency (F0) in the amplitude modulation of the pulse trains. Two nominal F0s representing male and female voices were included in the setup. The first SIPI approach inserted SIPIs at a rate equal to the F0 (full-rate SIPI, FRS) to support the F0 cue. The second approach inserted SIPIs at a rate one octave below the F0 (half-rate SIPI, HRS) to circumvent a potential rate limitation. Sensitivity was measured at five modulation depths (MDs, 0.1, 0.3, 0.5, 0.7, and 0.9) and with a carrier pulse rate of 2000 pulses per second. All signal conditions were individually loudness balanced. Our results for FRS show enhanced pitch sensitivity, especially at low MDs. HRS resulted in more ambiguous performance with listener-dependent benefit still being able to provide FRS-like benefit at higher F0s. In summary, the SIPI approach is likely to enhance both pitch and ITD sensitivity while maintaining speech intelligibility.

Zusammenfassung

Bei hochgradigem Hörverlust können Cochlea-Implantate (CIs) Teile der Hörwahrnehmung wiederherstellen, indem sie den Hörnerv direkt mit einem implantierten Elektrodenarray elektrisch stimulieren. Die zeitliche Feinstruktur der Audiosignale wird von den meisten Strategien zur CI-Signalkodierung verworfen. Daher wird Information zur Periodentonhöhe und zu interauralen Zeitdifferenzen (ITDs) nur unzureichend übertragen. Zusätzlich ist die Sensitivität für Periodentonhöhe und ITDs bei hohen Trägerpulsraten, die für die Sprachverständlichkeit wichtig sind, durch eine Ratenlimitierung beschränkt. Jüngst konnte mit einer neuen Kodierungsstrategie die CI-ITD-Sensitivität bei hohen Trägerpulsraten erhöht werden. Bei dieser Strategie werden Extrapulse mit kurzen Intervallen (SIPIs) periodischen Trägerpulsketten hinzugefügt. Im Rahmen dieser Masterarbeit wurde eine psychoakustische Studie mit sechs CI-Hörern durchgeführt, um den Einfluss zweier Ansätze zur Einfügung der SIPIs auf die Sensitivität für Periodentonhöhenunterscheidung zu untersuchen. Dabei wurden pseudo-syllabische Signale, welche stimmhafte Sprachsegmente durch die Kodierung der Grundfrequenz (F0) in der Amplitudenmodulation der Pulsketten imitieren, verwendet. Zwei nominale F0, die männliche und weibliche Stimme repräsentieren, wurden getestet. Beim ersten SIPI-Ansatz wurden Extrapulse mit einer Rate entsprechend der F0 eingefügt („full-rate SIPI“, FRS) und somit die F0 unterstützt. Beim zweiten SIPI-Ansatz wurden Extrapulse mit einer Rate eine Oktave unterhalb der F0 eingefügt („half-rate SIPI“, HRS), um den potentiellen Effekt einer Ratenlimitierung zu vermeiden. Die Sensitivität wurde bei fünf Modulationstiefen (MDs, 0,1, 0,3, 0,5, 0,7 und 0,9) und einer Trägerpulsrate von 2000 Pulsen pro Sekunde gemessen. Alle Signalbedingungen wurden lautheitsangeglichen. Die Ergebnisse für FRS zeigen eine Erhöhung der Sensitivität für Periodentonhöhe, insbesondere bei niedrigen MDs. Die Ergebnisse für HRS zeigen stark hörerabhängig positive oder negative Effekte. Dennoch zeigt bei höheren F0 auch HRS FRS-ähnlichen Nutzen. Die Verwendung von SIPI-Pulsen bei hohen Trägerpulsraten scheint damit eine Verbesserung der Periodentonhöhen- und ITD-Sensitivität unter Beibehaltung der Sprachverständlichkeit zu ermöglichen.

Affidavit

I declare that I have authored this thesis independently, that I have not used other than the declared sources/resources, and that I have explicitly indicated all material which has been quoted either literally or by content from the sources used. The text document uploaded to TUGRAZonline and KUGonline is identical to the present master's thesis.

Date

Signature

Acknowledgements

This thesis would not have been possible without the help and support of many people. I would like to express my heartfelt thanks to all of them. Some people deserve particular appreciation for their contributions:

First and foremost, I would like to thank Piotr Majdak for his excellent support, and for the many helpful tips regarding content and language.

I would like to thank Bernhard Laback for his outstanding support in the CI experiments and his decision to let me become part of this project for more than a year now.

I would like to thank Robert Höldrich for his help and the fruitful discussions in Graz.

Also, I would like to thank Sridhar Srinivasan for his help with the experiments and Piotr, Bernhard, and him for the extensive and fruitful discussions in Vienna.

I would like to thank all the ARI staff for allowing me to feel valued and part of the team from day one. Especially, I would like to thank Michael Mihocic for his help in finding the CI listeners and solving laboratory issues, and Maike Ferber for her help in collecting parts of the CI data.

Moreover, I would like to thank the CI listeners for their enthusiastic participation and patience with a total of almost 200 test hours.

Furthermore, I would like to thank my parents for supporting my studies for six years. Thanks to them I never had to worry about anything existential and could really enjoy the time and concentrate on my studies.

Finally, I would like to thank my girlfriend Anna for making my work in Vienna possible in the first place, for always patiently enduring my long monologues and for supporting me unconditionally.

Contents

List of Figures

List of Tables

1 Introduction

Hearing loss is one of the world's most common diseases. In the EU, 16 to 17 % of the population have a mild hearing impairment (HI) with better ear hearing levels (BEHLs) between 25 and 39 dB, 5 % have a moderate HI (BEHL 40 to 69 dB) and even 1 % (2017: approximately 5 million EU citizens [1]) suffer from severe to profound hearing loss (BEHL $\geq$ 70 dB). In 2025, 100 million EU citizens are estimated to have some kind of hearing impairment (Shield 2006). Recently, the WHO called for governmental action to cope with the ongoing rise in HI numbers (WHO 2018).

The consequences of hearing loss are manifold. Those affected often suffer from social isolation, a reduction of life quality, or even unemployment. In the EU, the estimated total costs per year have reached 224 billion €, with 35 billion € of these costs solely due to severe or profound hearing loss (Shield 2006). Hence, apart from improving the lives of hearing impaired people, there is also a huge financial interest of the societies and their social systems to reduce costs.

Because hearing loss is most often sensorineural, it affects all dimensions of auditory perception. For severe to profound hearing loss, cochlear implants (CIs) are normally the only treatment available. Despite the high costs for implantation and clinical treatment, CIs are increasingly implanted, resulting in a steadily growing population with electric hearing. Although substantial successes have been achieved in speech perception due to improved signal processing with strategies such as continuous interleaved sampling (CIS, Wilson et al. 1991), electric hearing is still far from providing normal hearing because of major deficits, e.g., in spatial hearing and pitch perception.

Within this thesis, I focus on pitch perception in electric hearing. First, I outline mechanisms of pitch perception in normal hearing and discuss CIs in terms of speech processing and pitch. Then, I motivate the psychoacoustical experiments forming the core of this thesis. Thereafter, I describe methods underlying the experiments, present the data collected, and analyze the statistically underpinned results. Finally, I discuss the results in the light of related literature, and draw general conclusions.

---

1. ec.europa.eu, accessed March 14, 2018

While working on the thesis, the data were presented to the scientific community as part of two conference talks:

(1) I have presented the data at 2018's Annual Conference of the "Deutsche Gesellschaft für Audiologie (DGA)":

Lindenbeck, M., Laback, B., Srinivasan, S. and Majdak, P. (2018), Enhancing rate-pitch sensitivity in electric hearing by inserting extra pulses with short inter-pulse intervals, *in* '21st Annual Conference of the German Society for Audiology, Halle (Saale), Germany', p. 131.

(2) Bernhard Laback used parts of the data I collected in his invited talk at the 175th Meeting of the Acoustical Society of America (ASA):

Laback, B., Srinivasan, S., Lindenbeck, M., Ferber, M. and Majdak, P. (2018), Towards increasing timing sensitivity in electric hearing, *in* '175th Meeting of the Acoustical Society of America'.

1.1 Pitch in Normal Hearing

The American National Standards Institute (ANSI) specifies pitch as "that attribute of auditory sensation in terms of which sounds can be ordered on a scale extending from low to high" (ANSI 1994). In music, sequences of pitch define melodies, simultaneous combinations form harmonies. In speech, pitch contours allow us to distinguish between statements and questions (prosody). Further, pitch helps us to identify the gender of speakers based on gender-specific ranges of the fundamental frequency (F0).

Pitch is duplex: Its acoustic basis is defined by the frequency on the one hand and by the period on the other. Though period (time domain) and frequency (spectral domain) are linked to each other, pitch perception would not be completely describable when throwing away one of the two quantities because their underlying mechanisms manifest at different auditory processing stages (e.g., Licklider 1951).

In human normal hearing, spectral analysis is performed by the basilar membrane (BM) in the cochlea which essentially performs a frequency-to-place mapping along the tonotopy. Because the tonotopic organization in the cochlea follows an approximately logarithmic scale from apex to base, frequency selectivity decreases with increasing frequency. The mapping is preserved at least up to the primary auditory cortex and forms the basis of the *place* pitch theory. Every tonotopic place of the cochlea is connected to neural populations that react purely to the temporal structure of the BM movement. The resulting neural firing patterns are the basis for the *temporal* or *rate* pitch theory.

Three different types of sounds have commonly been used to assess normal-hearing (NH) pitch perception: Pure tones, harmonic complex tones, and transposed tones.

Pure Tones. The pitch of pure tones is mostly characterized by their frequency which defines F0. In the cochlea, pure tones elicit a traveling wave (cf. Fig.1.1) that has its peak at the place corresponding to the tone's frequency, i.e., its tonotopic place.

The neurons of the auditory nerve (AN) transmit the temporal patterns of the BM movement. Here, the populations connected to that tonotopic place react most strongly. The pitch is coded using the so called phase locking property of AN fibers, i.e., their action potentials, or spikes, are locked to certain phases of the sinusoid periods. This information is related to as rate pitch or temporal pitch, respectively, and opens the possibility to extract periodicity information from neural firing by analyzing the inter-spike intervals (Oxenham 2013).

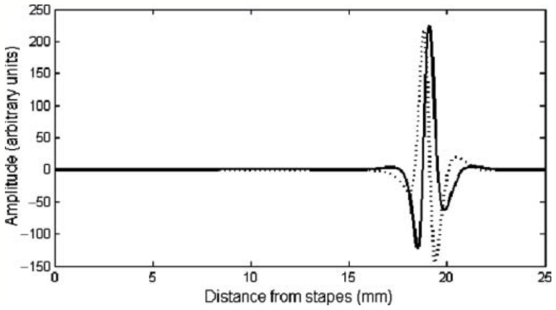

Figure 1.1 – Traveling wave on the BM at two time instances (Oxenham 2008).

Harmonic Complex Tones. This type of tones (e.g., saw-tooth signal shown in Fig. 1.2 A) consists of several harmonically related pure tones having frequencies that are integer multiples of the F0 (Fig. 1.2 B). These signals are much more common in nature (voiced speech, musical instruments, and so forth) than pure tones and they normally elicit a pitch corresponding to the F0, even when the signal does not contain the F0 ("residue pitch", Schouten 1940). These tones produce several excitation peaks on the BM which can be perceived as a single pitch.

The decreasing frequency selectivity with increasing signal frequency is often expressed in terms of auditory filters (Fig. 1.2 C), which model the cochlear filtering and broaden with increasing frequency: Each point alongside the BM only responds to a limited number of frequencies and can thus be modeled as a band-pass filter.

Hence, the excitation patterns on the BM (Fig. 1.2 D), which describe the response to a harmonic complex tone in the frequency domain, can be basically divided into two parts: The F0 and the low-numbered harmonics which produce distinct temporal neural structure, and the higher-numbered harmonics which produce smeared neural patterns. Thus, the lower-frequency parts of the signal that are processed individually at the outputs of the auditory filters are often referred to as *resolved*, the higher-frequency parts as *unresolved*.

In general, time domain signals can be factored into their rapidly varying fine structure (FS) and their slowly varying envelope (ENV) using, e.g., the Hilbert transform (Hilbert 1912). Smith et al. (2002) used this relation to create “chimaeric“ sounds in which the FS of a sound A is merged with the ENV of a sound B. They elegantly showed that ENV information is crucial for speech reception (at least in quiet) whereas the FS information is most important for pitch perception and low-frequency sound localization via interaural time

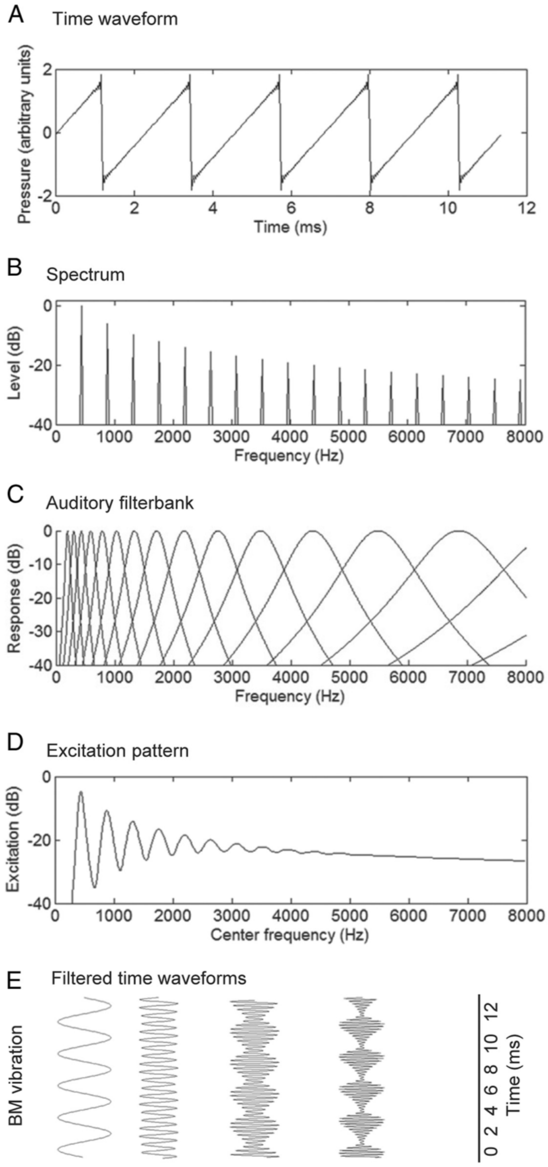

Figure 1.2 – Cochlear processing of a complex harmonic tone, F0 = 440 Hz (Oxenham 2012).

differences (ITDs).

After cochlear filtering, the resolved signal components are represented in the auditory periphery basically as separate pure tones. For these, the FS contains the temporal pitch in the form of the signal period, place pitch could be derived from the tonotopic excitation linked to the signal frequency. In contrast to that, the unresolved components form complex wave forms. Because based on F0, the ENV is always F0-periodic and thus theoretically contains the same temporal pitch information as the F0 itself (Fig. 1.2 E). Unlike pitch derived from FS cues, ENV-based pitch is much more prone to distortions between the phase relations of the harmonics, e.g., as a consequence of reverberations in the environment (Oxenham 2013).

Pitch sensations produced by resolved harmonics were found to be more salient and precise than those resulting from unresolved harmonics. This suggests that place or temporal FS information is more important for pitch perception. However, also temporal ENV information of the unresolved harmonics elicits a pitch sensation. Here the place cues are poorly encoded, leading to the conclusion that at least some of aspects of pitch can be extracted from purely temporal ENV cues (Oxenham 2008).

By using groups of frequency-modulated harmonic complexes, band-pass filtered in frequency regions differing in "resolvability", Carlyon and Shackleton (1994) hypothesized that the F0s of resolved and unresolved harmonics are processed by two separate mechanisms. Indeed, a one-dimensional model which combined F0 and filtered frequency region to define the resolvability of the harmonics was sufficient to predict the mechanism involved (Shackleton and Carlyon 1994).

Transposed Tones. Yet, solely time code is also not capable of explaining all aspects of pitch perception. Using transposed tones which map low-frequency resolved harmonics to high-frequency regions of the cochlea, Oxenham et al. (2004) found that complex tones with inconsistent place pitch information do not generate a residue pitch. This leads to the conclusion that, at least for harmonic complexes, high pitch sensitivity requires consistent place and temporal information. Oxenham (2008) suggested that in NH, all available information from both rate and place, or even mixtures of them (rate-place pitch), are used to increase the robustness of pitch estimations.

Neurophysiology. The psychophysical findings that two mechanisms are involved in human pitch perception are supported by neurophysiological data recorded in vocal primate species, i.e., awake macaque and marmoset monkeys. Steinschneider et al. (1998) suggested that in the primary auditory cortex, for F0s of 100 to 200 Hz (with an upper limit of 400 Hz), temporal features are encoded in basal cochlear regions by phase-locked neural activity, whereas place-related properties are encoded in apical cochlear regions for both F0 and harmonics based on the tonotopic excitation. Bendor et al. (2012) concluded that the pitch is encoded in ENV cues for lower-pitch sounds composed of unresolved harmonics and the pitch is encoded by place cues for higher-pitch sounds with resolved harmonics.

Eventually, spectral and temporal processing might be combined in the auditory cortex. This is supported by the existence of pitch-selective neurons in the primary auditory cortex, responding to both pure tones and complex sounds with a missing fundamental and the same F0 (Bendor and Wang 2005).

1.2 Cochlear Implants

CIs are currently the most widely used clinical treatment for severe hearing loss or deafness. Essentially, every implant, regardless of the manufacturer, shares the same four components (Fig. 1.3 A): (1) the external microphone and speech processor, (2) the external transmitter using electromagnetic induction to transmit the signals to (3) the internal receiver, and (4) the electrode array inside the cochlea. Further, for monopolar stimulation, a reference electrode located outside the cochlea is required.

The inserted n-channel electrode array optimally stimulates n tonotopic cochlear places (Fig. 1.3 B). Considering the rather low number of channels [12 (MED-EL) to 24 (Nurotron)], any place-dependent code is represented rather coarse in CI stimulation. Furthermore, depending on the distance from the implant to the AN and the individual AN sensitivities in general, the current elicited by an electrode spreads across a broad range of AN fibers. This results in channel interactions degrading the neural code (see, e.g., Wilson et al. 1991).

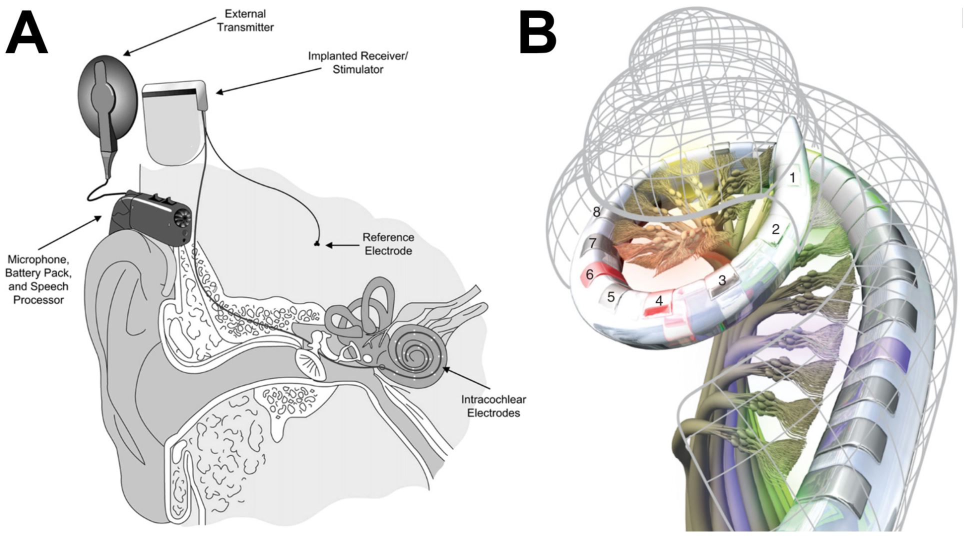

Figure 1.3 – A Components of modern CIs (Wilson and Dorman 2008) and **B** electrode array inserted into the cochlea. Each electrode is stimulating different fibers of the AN, spiral ganglion neurons, and the AN itself (Boulet et al. 2015).

1.2.1 Speech Processing

Despite lots of achievements, CIs still only restore parts of the normal auditory perception. The most important step in the development of sophisticated CIs was the invention of the continuous interleaved sampling (CIS) stimulation strategy (Wilson et al. 1991). Using multiple band-pass-filtered channels (Fig. 1.4) and across-channel interleaved ENV-based amplitude-modulated (AM) pulse trains, the CIS strategy first provided substantial speech recognition, even without visual cues. The main innovation was to rely on the listener's ability to extract cues from the channel signals instead of extracting specific perceptually relevant features of the acoustic signal using signal processing.

As pointed out by Smith et al. (2002), speech perception is based on both temporal ENV and spectral information across major parts of the cochlear tonotopy. Vowel recognition relies on the concentration of signal energy in certain frequency regions, the so called formants. More precisely, the formants F1 (frequency range 320 to 1000 Hz) and F2 (frequency range 800 to 3200 Hz) are essential. Consonants are broadband high-frequency sounds either with (voiced) or without (unvoiced) F0 information.

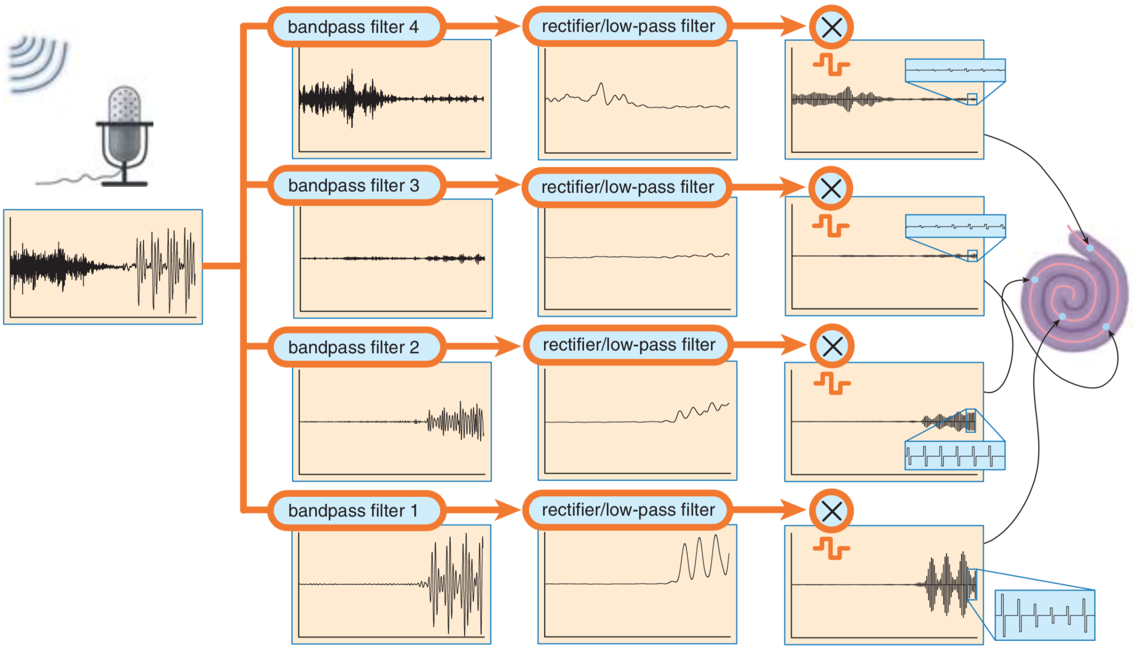

Figure 1.4 – CIS stimulation strategy: The input signal is a 100-millisecond slice of the syllable “sa“, containing the voiceless alveolar fricative “s“, the open front unrounded (voiced) vowel “a“, and the transition period. Exemplary, a four-channel (electrode) array is used (apex: channel 1; base: channel 4), dynamic range compression (automatic gain control, AGC) is omitted (Dorman and Wilson 2004).

As shown in Figure 1.5, already a low number of channels is sufficient to restore parts of speech perception in quiet in CI listeners. In context-free tasks (vowels/consonants), CI performance has been found generally worse than in normal hearing, even for the best CI listeners. The context itself provides some additional benefits resulting in less confusion of particularly similar sounds. In sentence recognition tasks, the star CI listeners reached NH performance (Friesen et al. 2001). In general, over the last 50 years sentence recognition scores in quiet improved to be satisfactory nowadays (Fig. 1.6).

Yet, some challenges remain (Fig. 1.7), e.g., tasks involving spatial hearing, pitch, tonal language processing, and also speech perception in noisy environments (Zeng et al. 2008). Many of them originate in structural deficits of the implants and the ENV-based CIS signal processing. In voiced speech as well as in background noise, FS cues and the F0 cue in particular become more important. Since the CIS stimulation strategy discards the acoustic FS by replacing it with periodic pulse train carriers, the F0 cue is only represented in the ENV (see Fig. 1.4 and the differences between “s“ and “a“) and is thus coded weaker than in normal hearing.

Speech recognition performance increases with increasing rates of the pulse train carriers. Loizou et al. (2000) suggested to use pulse rates above 800 pulses per second (pps) in order to improve consonant recognition. Similar results were also obtained by Arora et al. (2009), who measured CNC word recognition scores, and suggested pulse rates of at least 500 pps. In case low pulse rates are used, increase of the width of the carrier pulses also leads to improved consonant recognition (Loizou et al. 2000).

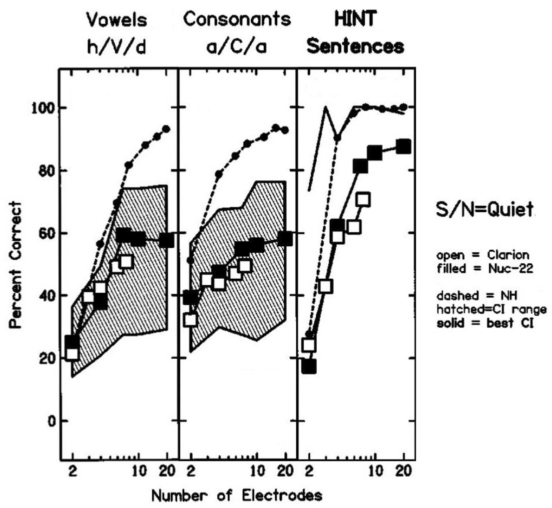

Figure 1.5 – Speech recognition in quiet as a function of spectral channels for medial vowels (left), medial consonants (middle), and HINT sentences (right) (Friesen et al. 2001).

1.2.2 Pitch

The properties of pitch perception in electric hearing differ substantially from those in normal hearing. The CIS-like stimulation strategies focus on the transmission of speech information, thus extract the ENV information of different acoustic frequency bands, roughly matched to the electrode positions in the cochlea. Place information is limited by the number of electrodes, the temporal FS is discarded completely.

Still, using direct stimulation by means of a research interface which bypasses the speech processors, CI listeners are able to perceive both temporal pitch, e.g. by varying the rate of unmodulated low-rate pulse trains at one electrode (see, e.g., McDermott and McKay 1997), and place pitch, e g., by keeping the pulse rate fixed and stimulating different electrodes (see, e.g., McDermott and McKay 1997). The listeners’ ability to discriminate temporal pitch is limited to approximately 300 pps though (Shannon 1983, Townshend et al. 1987, Ihlefeld et al. 2015). The neural mechanism underlying this lim-

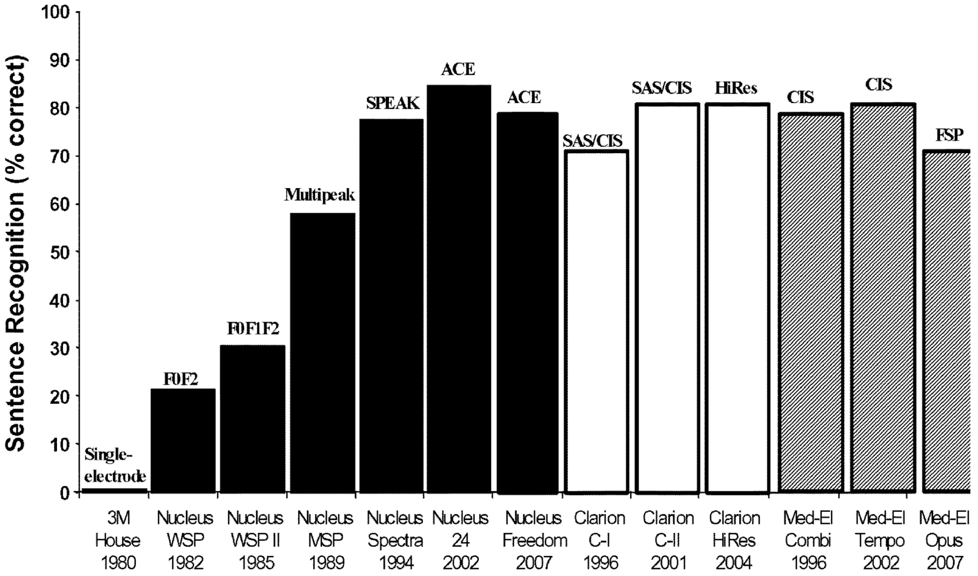

Figure 1.6 – CI sentence recognition scores in quiet, different manufacturers, implant models, and stimulation strategies (Zeng et al. 2008).

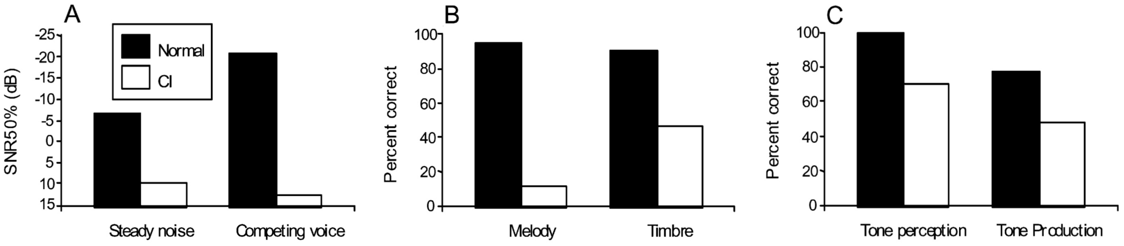

Figure 1.7 – Performances of CI listeners in challenging tasks: *(A)* speech recognition in noise[2], *(B)* music perception and *(C)* tonal language processing (Zeng et al. 2008).

itation remains unknown as mammalian phase locking is accurate up to rates of 1 kHz in both normal hearing (Rose et al. 1968) and electric hearing (Dynes and Delgutte 1992). This limitation can be partially explained by a saturation of neural discharge rates due to neural refraction (for a summary of neural stimulus-response phenomena associated with cochlear implants see Boulet et al. 2015). Above 300 pps, place pitch is more dominant, but the cue is less salient and more influenced by changes in loudness (Pijl 1997, Zeng 2002). Following the results from Arnoldner et al. (2008), roughly 87 % of the CI listeners perceive changes in pitch based on variations in loudness. For 73 % of these lis-

---

[2] “SNR50%“ describes the SNR needed to achieve 50 % correct responses.

teners, pitch decreased with increasing intensity whereas it increased for the other 27 %. This effect was consistent across cochlear locations. A similar trend was reported by Vandali et al. (2013) who found that pitch decreased with increasing loudness for half of the listeners. Further, CI listeners' rate pitch discrimination performance has been found to be worse than that in normal hearing, reflected by the poorer rate discrimination just-noticeable differences (JNDs) of roughly a semitone (McDermott 2004, Oxenham 2008, Kreft et al. 2010, Green et al. 2012, Vandali et al. 2013).

In a CIS-like stimulation, periodicity is essentially transmitted via AM only. McKay et al. (1994) demonstrated, that robust pitch perception based on AM is possible only if the AM frequency is sufficiently oversampled by the carrier (i.e., an oversampling factor of four to five, see also Wilson et al. 1997). McKay et al. (1995) refined this outcome by suggesting that the pitch perceived with AM stimuli can be predicted by an average of the carrier and the modulation rate, exponentially weighted with the modulation depth (MD). Here, starting at the carrier rate, with increasing MD, the pitch approached the modulation rate as well as the pulse repetition rate of unmodulated pulse trains[3] (Vandali et al. 2013). Since the AM contains acoustic ENV information, McKay and Carlyon (1999) suggested that ENV pitch in EH is comparable to the residue pitch of high-frequency band-pass-filtered (3.9 to 5.3 kHz) acoustic pulse trains in normal hearing, thus not containing resolved harmonics. By comprehensively testing a "star" CI listener, (McDermott and McKay 1997) showed that musical pitch information can be conveyed within roughly two octaves. They varied the pulse repetition rate, the modulation frequency of AM pulse trains, the place of stimulation, or combinations of rate and place. When both rate and place were varied, the place pitch dominated.

The psychophysical results supporting ENV-based pitch perception are further confirmed by neurophysiological data showing that most neurons strongly phase lock to the peak of suprathreshold ENV phases. The extent of phase locking is higher than in normal hearing which thus eliminates spike timing as a cue to the ENV shape (Hancock et al. 2017). As CI signals do not contain frequency modulation (FM) of the carrier pulse trains, the weakness of pitch perception in electric hearing might also be explained by missing FM. However, Brochier et al. (2018) suggested no difference between AM and FM-based pitch perception.

---

[3] The pulse repetition rate was equal to the modulation rate.

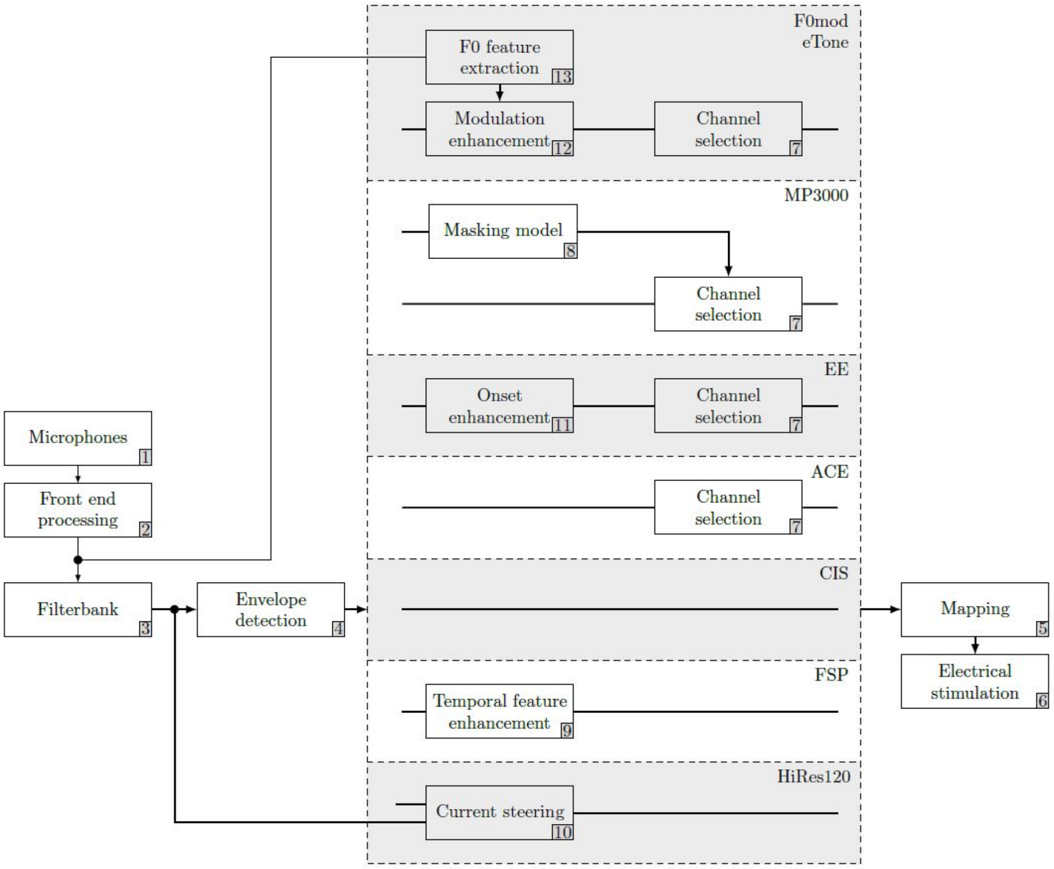

Figure 1.8 – Examples of modified monaural sound processing strategies from Wouters et al. (2015). Common parts and modifications in order to improve the transmission of temporal cues.

As pitch perception is largely affected by current stimulation strategies succeeding in speech perception, many attempts were made to include more salient temporal cues in the signals using modified versions of the classic (monaural) CIS approach (Fig. 1.8, Wouters et al. 2015). In general, all of these strategies still rely on filter-bank analysis and per-band ENV extraction. Modifications could be grouped in two clusters: First, many strategies apply an “n-of-m“ channel selection approach and stimulate only the most informative channels (e.g., ACE). Doing so reduces channel interactions and precises the transmitted temporal patterns. Second, temporal features such as the F0 are extracted and artificially emphasized by, e.g., enhancing the AM (F0mod/eTone) or encoding temporal fine structure from low rates at apical electrodes (FSP).

Other researchers also investigated, how place-pitch perception might be improved. Exemplarily, proposed approaches modified the filter bank, e.g., by ensuring that the first harmonic is always resolved in two adjacent filters (Geurts and Wouters 2004) or used asymmetric pulses (see, e.g., Macherey and Carlyon 2012). However, all these concepts cannot overcome the limitations induced by the technological restrictions. By using vocoder techniques, Mehta and Oxenham (2017) found that at least 32 electrodes with no channel interaction or 64 electrodes with filter slopes of more than 72 dB per octave are needed for complex pitch perception when ruling out ENV and spectral edge pitch cues [4].

Based on the findings presented in this chapter, the motivation for our experiments is summarized in the following chapter. The new signal processing approach that is used aiming at improving temporal pitch is explained. Further, research questions are developed and the setup of our experiments is outlined.

---

[4] Spectral edge pitch cues produce a pitch sensation associated with complex tones' extreme spectral components.

2 Motivation

Within the psychophysical study, we focus on improvements in *temporal* pitch perception, as place pitch is largely affected by hardware restrictions of current electrode arrays. Improving the temporal coding in CI signals might thus be beneficial for both current as well as future implantees with possibly modernized electrode arrays.

2.1 Context: Short-Interpulse-Interval (SIPI) Stimulation

Laback and Majdak (2008) were able to restore ITD sensitivity at pulse rates above 300 Hz where sensitivity is normally absent in electric hearing (e.g., Laback et al. 2007, van Hoesel 2008, van Hoesel et al. 2009, Ihlefeld et al. 2015). To accomplish that, they jittered the interpulse intervals (IPIs) of CI pulse trains in a binaurally coherent way[5] (Fig. 2.1). Neurophysiological data by Hancock et al. (2012) revealed that sensitivity was restored due to the jitter randomly creating very short IPIs (SIPIs). For a more detailed summary of jitter-related research, see Laback (2012).

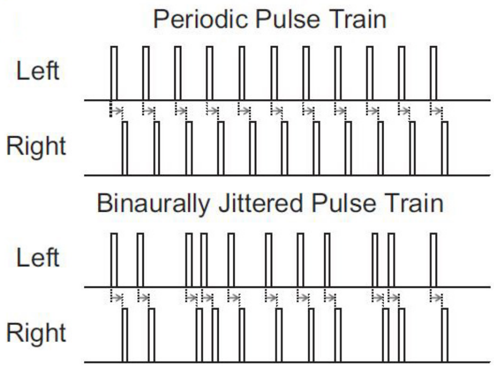

Figure 2.1 – Periodic pulse trains (upper) and jittered pulse trains (lower) preserving ITD sensitivity at high rates, depicted for both left and right ear (Laback and Majdak 2008). The arrows indicate the ITDs coded in the pulse trains.

The finding that SIPIs are the basis of the increased ITD sensitivity with jittered pulse trains is especially interesting for both ITD and rate pitch perception, because it would allow the pulse trains carriers of ENV-based high-rate stimulation strategies to be modified in a deterministic way, e.g., by introducing SIPIs at fixed repetition rates, thus exploiting the benefits of the SIPIs while excluding detrimental side effects of the stochastic jitter approach such as randomly sampled speech ENVs and thus weakened speech understanding and F0 perception.

---

[5] Binaurally coherent jitter preserves the ITD because it is applied identically to both ears.

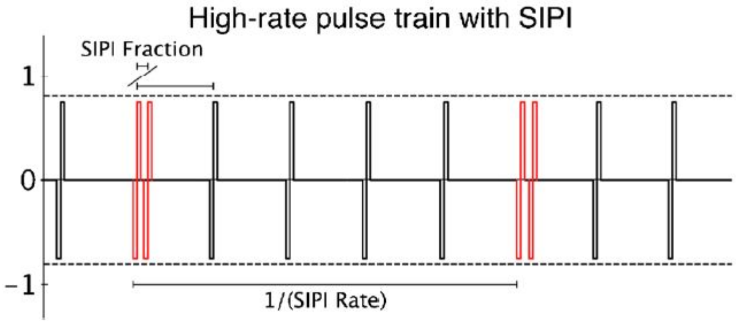

Figure 2.2 – SIPI parameters “fraction“ and “rate“ (Srinivasan et al. 2018).

Srinivasan et al. (2018) systematically investigated ITD sensitivity with SIPI stimulation using “laboratory signals“, i.e., unmodulated 1000-pps pulse trains. They introduced two basic parameters characterizing the SIPI insertion (Fig. 2.2): The “SIPI rate“ describes the rate at which SIPI pulses are inserted into the carrier pulse train. The “SIPI fraction“ depicts the IPI between a carrier pulse and the subsequent SIPI pulse relative to the carrier IPI. Note, that in case of a constant SIPI fraction, the SIPI rate has to be an integer sub-multiple of the carrier pulse rate. The authors concluded that introducing SIPIs with rates $\leq$ 100 pps and fractions $\leq$ 20 % substantially enhances ITD perception.

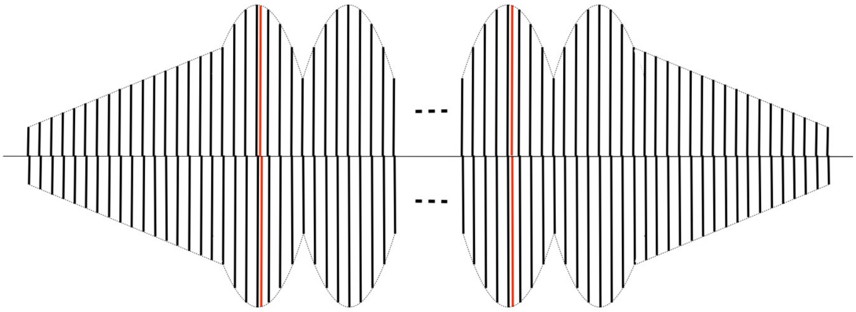

Figure 2.3 – Pseudo-syllabic stimulus: AM high-rate pulse train containing SIPIs in *red* (Srinivasan et al. 2017).

In order to also investigate the effect of SIPI pulses in signals closer to those produced by CIS-like stimulation strategies, Srinivasan et al. (2017) collected data on ITD discrimination sensitivity when using so-called “pseudo-syllabic“ stimuli (Fig. 2.3). Essentially, the signals were periodical AM signals where the AM rate was equal to the F0 such as in voiced speech segments. For more details on the signals, see Lindenbeck (2017). The authors kept the SIPI rate at 62.5 pps and used two AM rates of 125 and 250 Hz, respectively, as well as five different MDs (0.1, 0.3, 0.5, 0.7, and 0.9). The data showed improvements for low to moderate MDs for both AM rates (125 Hz: 0.1 to 0.5; 250 Hz: 0.1 to 0.3) when SIPIs were at AM phases near the peak. Considering the typically low MDs in everyday CIS signals (mean MD 0.25, range 0.1 to 0.4[6]), SIPI pulses might also be beneficial in “real-life“ situations.

There is also another potential advantage of the SIPI approach compared previous approaches which focused on the enhancement of the envelope in order to better code pitch by, e.g., artificial enhancement of F0 fluctuations in the envelope (Geurts and Wouters 2001) or F0-amplitude modulating the channels of an n-of m (ACE, cf. Fig. 1.8) stimulation strategy with 100 % MD (Laneau et al. 2006, Milczynski et al. 2009). While psychophysical data showed improvements for the latter approach, the impact on the “natural“ envelope containing the speech-relevant cues was substantial, though. Contrary to that, the SIPI approach preserves the envelope.

2.2 Research Questions

The main research question is to investigate the influence of the SIPI approach on temporal pitch perception with pseudo-syllable signals. If there is an effect, we aim to investigate the following details:

(1) How does the efficiency of the SIPI approach depend on the MD?

(2) Is pitch perception with SIPIs comparable to that with (unmodulated) low-rate (LR) pulse trains?

(3) Does pitch perception change with F0, possibly indicating a rate limitation? If so, does lowering the SIPI rate (relative to the F0) allow to circumvent that rate limitation?

(4) Does pitch perception change with carrier rate, having influence on applicability?

---

[6] Proposal of National Institutes of Health (NIH) project “Bilateral Cochlear Implants: Physiology and Psychophysics“, grant #: R01DC005775.

2.3 Outline of the Experiments

The psychoacoustical study is structured into two major parts, the pretests and the main experiment.

Pretests. While we focus on a pitch discrimination task, loudness changes covarying with the pitch confound the statistical outcomes of temporal pitch studies (e.g., Carlyon et al. 2010, Chatterjee and Oberzut 2011, Vandali et al. 2013). To avoid loudness differences between conditions, we first loudness balanced all experimental conditions before starting the actual pitch-related experiments. Afterwards, all listeners were trained on the pitch discrimination tasks in order to rule out task learning effects. In the pitch discrimination pretest, we then measured the individual pitch discrimination sensitivity in order to adjust the parameters for the main experiment.

Our pitch discrimination experiments used two intervals in each trial. Each of the intervals contained a different F0, and the difference between the two F0s was the listener-specific frequency difference (FD).

Main Experiment. We measured pitch discrimination sensitivity without SIPIs (NS condition) as the reference and with SIPIs in the SIPI condition. Two differently motivated configurations were used:

(1) *SIPI at full rate (FRS):* The F0 was coded in the ENV by adding SIPIs at a rate of the F0. Relative to the F0, the SIPI pulses are thus inserted at “full rate“, meaning a single SIPI pulse is inserted within every AM period.

(2) *SIPI at half rate (HRS):* As F0s roughly cover a frequency range from 80 to 300 Hz with mean F0s of 120 Hz for male and 210 Hz for female speakers[7], high F0 cues might already be altered by the previously discussed rate limitation. Further, ITD data (Srinivasan et al. 2018) suggest that SIPI provides most improvements below 100 pps. To consider this, we also investigated the idea of inserting SIPI pulses in every other AM period. Relative to the F0, the SIPI pulses were thus inserted at “half rate“.

---

[7] Proposal of National Institutes of Health (NIH) project “Bilateral Cochlear Implants: Physiology and Psychophysics“, grant #: R01DC005775.

MD is a key factor to the salience and the accuracy of the F0 pitch cues. Hence, we varied the MD across its entire range (0.1, 0.3, 0.5, 0.7, and 0.9). Contrary to that, many laboratory studies only used rather high modulation depths [e.g., 80 % (Landsberger 2008) or 100 % (Kreft et al. 2010, Galvin et al. 2015)]. However, in everyday situations the MD is often reduced.

In normal hearing, pitch is an essential cue to gender which can be useful in challenging listening situations to segregate sound sources. To cover the F0 range across genders, we tested the pitch discrimination sensitivity at two nominal F0s namely low (125 Hz) F0 representing male speakers and high (250 Hz) F0 representing female talkers.

The actual F0s in our experiments exceeded 250 Hz and even reached 300 Hz. To still sufficiently oversample the ENV based on the "sampling theorem for electric hearing“ (McKay et al. 1994, Wilson et al. 1997), the main carrier frequency in the study was fixed at a relatively high rate of 2000 pps. Using 2000 pps allowed us to test sensitive participants at a FD of a semitone which is often set as the temporal pitch discrimination JND in electric hearing (see Sec. 1.2.2). However, carrier rate of 1000 pps was also used to test a carrier rate more typical for clinical CI systems and to better compare pitch and ITD data.

Because low-rate pulse trains provide an excellent pitch perception, we also measured the discrimination performance at three low rates, namely 62.5 pps, 125 pps, and 250 pps, representing all possible SIPI rates arising from combinations of nominal F0 and SIPI condition.

In the following chapter, the setup and the results of the pretests are presented. The results are discussed either as proof of concept or in context with the setup of the main experiment.

3 Pretests

Before conducting the main experiment some pretests had to be made to rule out confounding loudness cues, train the listeners on the task, and match the level of difficulty in the main experiment to the listener's individual discrimination sensitivity.

3.1 Methods

Participants. Six adult listeners [four females, two males; five post-lingually deafened, one pre-lingually deafened (CI21)] participated in the study. All had uni- or bilateral 12-channel MED-EL implants (more details are listed in Tab. 3.1). Bilateral participants (all except CI18) were asked to choose their preferred ear. Except for CI21, the decisions corresponded with the ear that was implanted first.They were paid an hourly wage for their participation and all procedures involving human subjects were authorized by the ethics board of the Medical University of Vienna (vote #2155/2013).

Listener	**Etiology**	**Age at testing (yr)**	**Age at onset of deafness (yr)**	**Ear used for testing**		
				Side	CI training (yr)	El #
CI17	Idiopathic	71	41	R	13	8
CI18	Sudden hearing loss	60	adult	R	13	8
CI21	Congenital	24	0	R	19	8
CI24	Progressive	55	40	L	13	8
CI74	Sudden hearing loss	48	41	R	6	8
CI77	Sudden hearing loss	64	58	R	6	8
Median	—	*58*	*41*	—	*13*	*8*

Table 3.1 – Details on the CI listeners taking part in the experiments.

Apparatus. The implants were connected to a personal computer via the Research Interface Box II (RIB2, Institute of Ion Physics and Applied Physics, Leopold-Franzens-University of Innsbruck, Austria), thus bypassing the clinical signal processors and making a soundproof chamber superfluous. Nevertheless, participants were protected against irritating visual cues. A customized version of the ExpSuite software (Lindenbeck 2017) was used to control the experiments, i.e., generate the stimuli, play them to the CI listener, and protocol the responses.

Stimuli. 600-ms pseudo-syllabic signals were used (Lindenbeck 2017, Srinivasan et al. 2017). The carrier rates were 2000 pps and 1000 pps. To avoid onset and offset cues, 150-ms ramps were applied. The ramps never contained SIPIs. The transitions between the ramps and the steady state part were ensured to be smooth. Further, the length of the steady state was adjusted to contain an integer number of periods corresponding to the F0s. To avoid F0-specific length cues, this number of periods was roved by ± 1. Previous studies (Srinivasan et al. 2017, Hu et al. 2017) revealed that explicit coding of ITDs might be most beneficial at the AM peak in electric hearing. Further, preliminary data on rate pitch (Lindenbeck 2017) did not suggest a difference between pulse insertions at the onset and the peak. Thus, SIPIs were inserted at peaks of the AM and it was ensured that there was a carrier pulse at the peak of the ENV. The SIPI fraction was restricted (at its lower boundary) by the pulse characteristics and was set to 12 % (2000 pps carrier rate) and 6 % (1000 pps carrier rate) resulting in the same time gap between carrier and SIPI pulse.

The carrier pulse trains consisted of biphasic pulses with a phase duration of 26.7 μs and a minimum inter-pulse gap of 1.7 μs yielding a pulse duration of 55.1 μs. The implants were used in monopolar mode. The stimuli were presented to a single electrode. According to (Baumann and Nobbe 2004), electrode eight is ideally located close to the tonotopic location of the carrier rate in normal hearing. Thus it was selected in order to closely match rate and place pitch cues (for data from normal hearing, see Oxenham et al. 2004).

Task. Listeners were tested using a 2I-2AFC[8] tasks. They either had to indicate which stimulus in the two intervals was louder or whether the stimulus in the second interval was

---

[8] two-interval two-alternative forced-choice task

higher or lower in pitch than the stimulus in the first interval. Participants provided their responses via a hand-held controller. They were given feedback to reduce response bias (Klein 2001). Before each new task, listeners were orally instructed. Written instructions were available at all time.

In each trial, the intervals were separated by an inter-stimulus gap of 490 ms. After each response, a new trial was played following a 310 ms break. If a listener felt uncomfortable with the stimulus timing, the inter-stimulus gap was adjusted.

Fitting. Two fittings were made with unmodulated pulse trains (without SIPIs) in order to determine the individual threshold (THR), comfortable level (CL), and maximum comfortable level (MCL). The first was conducted with 2000 pps used for the pseudo-syllabic signals, and the second with 125 pps used for the LR signals.

The general objective was to use electrode eight for all participants. Hence, the fitting was made for this electrode and the dynamic range (DR) was evaluated. Since the DR was sufficiently large for all listeners, they were all tested at electrode eight (cf. Tab. 3.1).

3.2 Loudness Balancing

Procedure. We measured the electric current needed for balanced loudness across conditions using an adaptive 3-up 1-down staircase procedure (cf. Jesteadt 1980). Staircases were randomly interleaved across upward and downward (inverse decision rules) staircases, each staircase was repeated[9] twice. Staircases terminated after reaching twelve turnarounds. The loudness-balanced currents were calculated by averaging the last eight turnarounds and subsequently averaging all results for one condition.

The reference was the 2000-pps, 125 Hz-F0, 0.5-MD, NS signal condition. Target stimuli depended on the carrier rate. For 2000 pps, targets were all combinations of MD (0.1, 0.3, 0.5, 0.7, 0.9), F0 (125 and 250 Hz), and SIPI (NS, FRS and HRS). Note that the NS condition was loudness balanced for an F0 of 125 Hz only because we assumed that changes in F0 do not evoke changes in loudness (Chatterjee and Oberzut 2011, Vandali et al. 2013). For 1000 pps, target was a 125-Hz F0, 0.5-MD, NS stimulus and it was balanced against the 2000-pps reference. The amplitudes for other 1000-pps conditions

---

[9] Note that the term “repetition“ here refers to the total amount of runs.

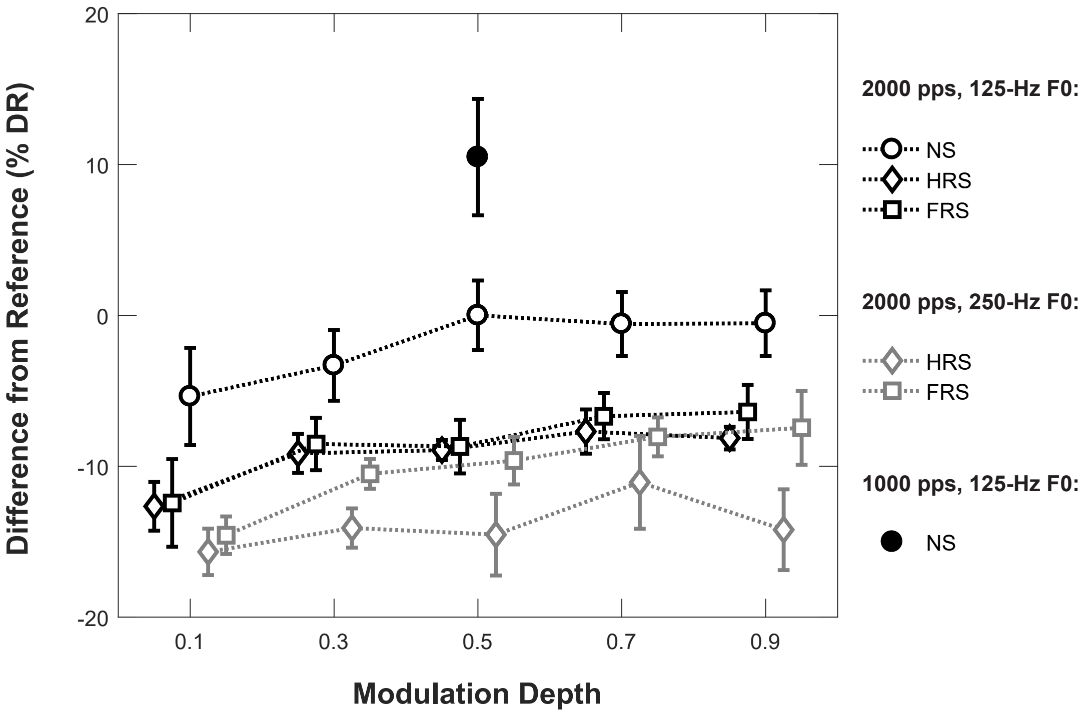

Figure 3.1 – Peak amplitudes after loudness balancing for the pseudo-syllable signals, averaged across listeners. Amplitudes adjusted softer than the reference indicate higher loudness of the corresponding signal condition. Error bars show normalized standard errors of the mean (SEMs, Cousineau 2005), corrected after Morey (2008).

were extrapolated based on the 2000-pps amplitudes. In order to evaluate the extrapolation, participants were played selected 1000-pps conditions and informally asked to indicate if both reference and target were equally loud. If necessary, the amplitude of the target was adjusted.

Results and Discussion. Since the loudness balancing was time consuming, we briefly checked out of curiosity whether this extensive setup was indeed necessary retrospectively, i.e., we investigated the influence of MD, SIPI, F0, and carrier rate on the peak amplitudes. Thus, in Figure 3.1, the high-rate conditions are displayed as differences of the peak amplitudes after loudness balancing to the reference in % DR.

As for MD and SIPI, a two-way repeated measures (RM) analysis of variance (ANOVA) with factors MD (0.1, 0.3, 0.5, 0.7, and 0.9) and SIPI (NS, FRS, and HRS) was performed on the 2000-pps data set. Both effects were significant[10], MD [$F(4,114) = 7.221$, $p < 0.0001$] and SIPI [$F(2,114) = 48.20$, $p < 0.0001$]. Post-hoc Tukey-HSD tests[11] revealed significantly lower peak amplitudes for MD of 0.1 compared to MD ≥ 0.5. Visual inspection of the individual data (not shown) indicated that the data from CI77 differ from the rest of the participants. Thus, the two-way RM-ANOVA with factors MD and SIPI condition was repeated while excluding the data from CI77. The outcome was similar apart from the effect of MD not being significant [$F(4,90) = 2.19$, $p = 0.0763$]. Post-hoc tests showed no significant differences between MDs. The effect of SIPI condition was still significant [$F(2,90) = 69.61$, $p < 0.0001$].

As for the F0, a one-way RM-ANOVA was performed on the 2000-pps data set excluding NS conditions. It showed a significant effect of F0 [$F(1,94) = 5.54$, $p < 0.05$] with significantly lower amplitudes for the 250-Hz F0.

As for the carrier rate, a paired t-test was run between the reference condition and the target condition only differing in the carrier rate. This test just failed to reach significance [$t(4) = 2.71$, $p = 0.0532$].

The statistical analysis revealed significant influences of three out of four parameters considered in the setup of the loudness balancing (MD, SIPI, and F0). This suggests that including conditions from all combinations of these parameters in the setup was necessary. The influence of the carrier rate did just not reach significance. However, because of the tight result, it might be safer to also consider the carrier rate.

The results seem to be in line with a loudness model by McKay and Henshall (2009) stating that, for a given carrier rate, the peak amplitude is the key factor for the loudness of AM signals. In our study, this is reflected in five of the six participants.

---

[10] The significance level was 5 %.

[11] Until further notice, all post-hoc tests mentioned are of this kind.

3.3 Pitch Discrimination

Before starting the main experiments, participants were trained on the pitch discrimination task. Subsequently, their individual pitch discrimination sensitivity was estimated to adjust the parameters of the main experiment.

Procedure. The tests used the method of constant stimuli. Each condition was repeated 100 times with randomized but balanced order of intervals. Further, trials were randomized across conditions. Performance is denoted in d' (Klein 2001), accounting for response bias and improving variance homogeneity compared to a percent scale. A d' of 1 is equivalent to a bias-free score of 76 % correct responses. The nominal F0 was the geometric mean of upper and lower F0 (cf. Kreft et al. 2010) which allowed to better center the tested F0s around the nominal F0. 3 %-DR level roving was applied independently across the two intervals of a trial to avoid potentially confounding loudness cues covarying with F0. Only 2000-pps signals were used and all conditions were tested together in randomized order. Participants were able to make breaks if they wanted so.

3.3.1 Training

Nominal F0 (Hz)	FD (%)	Lower F0 (Hz)	Upper F0 (Hz)	Geometric Mean F0 (Hz)
low	67	100	167	129
	100	91	182	129
high	83	182	333	246

Table 3.2 – Nominal F0s and FDs used in the training.

All participants were trained on the pitch discrimination task using frequencies that are assumed to be easy to discriminate (Tab. 3.2). To achieve that stimuli with large FDs, a MD of 0.3, and NS or FRS were used.

In the FRS conditions, all participants had a $d' \geq 1$ (data not shown). In the NS conditions, all but CI24 had a $d' \geq 1$.

3.3.2 Pitch Discrimination Pretest

In order to select a constant FD which avoids floor and ceiling effects, we pretested participants' pitch discrimination sensitivity.

Conditions. For each nominal F0, we used three conditions, MD 0.1/NS, MD 0.1/FRS, and MD 0.7/FRS, and three FDs. Details are denoted in Table 3.3.

Nominal F0 (Hz)	FD (%)	Lower F0 (Hz)	Upper F0 (Hz)	Geometric Mean F0 (Hz)
low	6	118	125	121
	20	111	133	122
	46	105	154	127
high	13	222	250	236
	29	222	286	252
	57	182	286	228

Table 3.3 – Nominal F0s and FDs used in the pitch discrimination pretest.

The MD 0.1/NS condition was assumed to be most challenging, the MD 0.7/FRS condition was assumed to be least challenging. The three FDs were chosen considering:

(1) the smallest FD possible [12] with the experiment setup. Thus, participants being "too" sensitive and showing ceiling effects already for the lowest FD would have been excluded from the main experiment;

(2) all FDs were smaller than those used in the training;

(3) at least one intermediate FD between the FDs used in the pretest could be realized in the main experiment providing us more options to decide.

The individual FDs were chosen such that the performance difference between the MD 0.1/NS and at least one of the other conditions was maximal. Further, sufficient space to

---

[12] Due to the 2000-pps carrier rate, the lowest FDs were 6 % (roughly 1 semitone) for the low nominal F0 and 13 % for the high nominal F0. This is in line with literature discussed in Section 1.2.2 stating that the average CI pitch discrimination JND ranges roughly at 6 %.

the floor ($d' = 0$) and the ceiling ($d' = 3.3$ [13]) of the d' score was ensured. d's close to the floor for all conditions across the tested FD range indicated (too) poor sensitivity for the main experiment. Opposite to that, d's close to the ceiling for all conditions across the entire FD range indicated (too) high sensitivity.

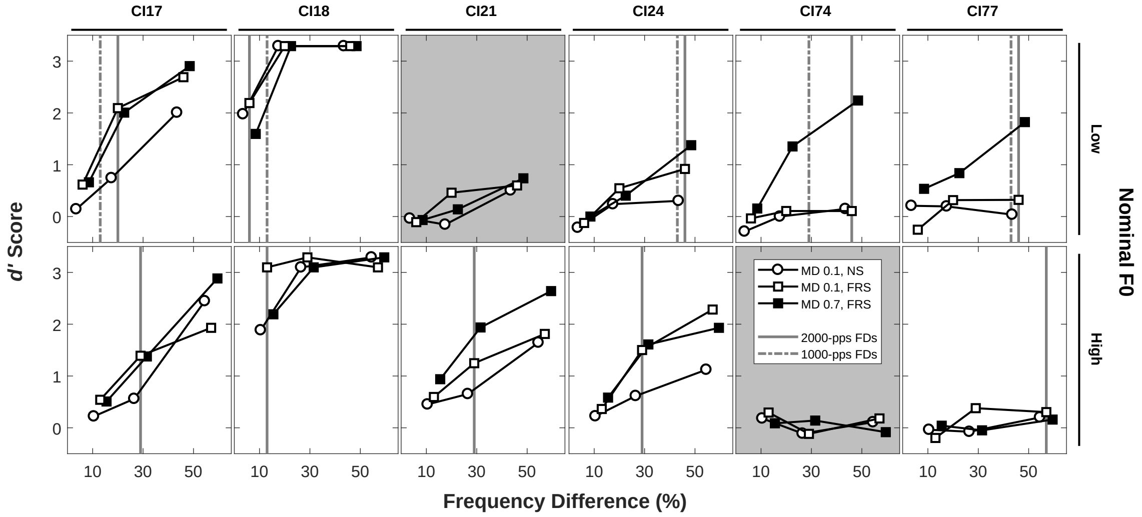

Figure 3.2 – Individual pretest d's: Both nominal F0s, three stimulus conditions. The chosen FDs for all combinations of nominal F0 and carrier rate used in the pitch discrimination experiment are indicated as vertical lines. Conditions that were not tested in the pitch discrimination experiment are shaded gray.

Results. The individual results are shown in Figure 3.2 as a function of FD separately for both nominal F0s. The results indicate no sensitivity for CI21 at the low nominal F0 and for CI74 at the high nominal F0. The results also show that CI18 was an exceptional performer indicating high sensitivity down to the lowest possible FDs. Given that different carrier rates result in different realizable FDs, decisions for both carrier rates used in main experiment (2000 pps and 1000 pps) were made based on the 2000-pps results for the low nominal F0.

Figure 3.3 shows d's averaged across listeners. Further, the average FD across the listeners and its standard deviation (SD) is shown. In general, the average FDs were quite similar across the conditions, but note the large variance. When considering the low

---

[13] A d' score of 3.3 is equivalent to a bias-free 99 %-correct response performance. The maximum percent-correct performance is set to 99 % to prevent the d' score from reaching infinity.

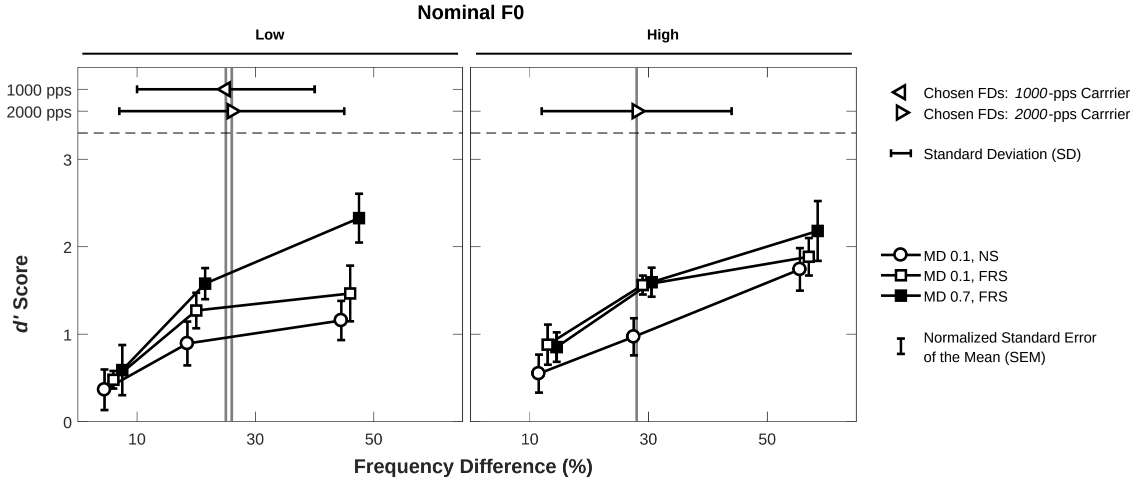

Figure 3.3 – Pretest d's averaged across listeners. The conditions highlighted in gray in Fig. 3.2 were excluded. The vertical error bars indicate normalized SEMs. The horizontal error bars denote the SD of the listener-specific FDs chosen for the pitch discrimination experiment. The dashed horizontal lines correspond to the maximum d' score.

nominal F0 only, the SD was smaller for the 1000-pps carrier compared to the 2000-pps carrier. This reflects the limited number of realizable FDs (the lowest FD for the 1000-pps carrier is 13 %, as used for CI18).

At the high nominal F0, the average FD nicely matched the point on the FD scale where performance difference between the toughest and the easiest signal condition (MD 0.1/NS and MD 0.7/FRS) was maximal. This opened the possibility to search for differences across signal conditions in the main experiment. At the low nominal F0, the performance difference between the easiest and the toughest condition monotonically increased with increasing FD. For the average FD, the difference between the d's of the two conditions is even slightly higher than at the high nominal F0. This suggests that also for the low nominal F0, inspections of differences across conditions in the main experiment are promising. Finally, the averaged d's also suggest sufficient room to the floor and ceiling of the d' scale.

Each combination of nominal F0, FD, and carrier rate resulted in a unique combination of upper and lower F0 and thus their geometric mean (GM). The GM was aimed to be as close as possible to the nominal F0. Table 3.4 contains the individual parameters per participant as well as the group statistic. Although the GMs of the interval F0s slightly

Listener	**Carrier Rate: 2000 pps**								**Carrier Rate: 1000 pps**			
	Low Nominal F0				**High Nominal F0**				**Low Nominal F0**			
	FD	**L**	**U**	**GM**	**FD**	**L**	**U**	**GM**	**FD**	**L**	**U**	**GM**
CI17	20	111	133	122	29	222	286	252	13	111	125	118
CI18	6	118	125	121	13	222	250	236	13	111	125	118
CI21	—	—	—	—	29	222	286	252	—	—	—	—
CI24	46	105	154	127	29	222	286	252	43	100	143	120
CI74	46	105	154	127	—	—	—	—	29	111	143	126
CI77	46	105	154	127	57	182	286	228	43	100	143	120
GM	*26*	*109*	*143*	*125*	*28*	*213*	*278*	*244*	*25*	*106*	*136*	*120*
SD	*19*	*6*	*14*	*3*	*16*	*18*	*16*	*11*	*15*	*6*	*10*	*3*

Table 3.4 – Lister-specific setup of the pitch discrimination experiment for all three tested combinations of carrier rate and nominal F0. Per combination: FD in %, and lower F0 (L), upper F0 (U), and their geometric mean (GM) in Hz; GM and SD for the sample. No data for one combination indicate that the participant did not show sensitivity for the nominal F0 involved.

varied from the intended nominal F0, the group means reached the nominal F0s nicely. Further, average FDs and their SDs are similar across nominal F0s.

After having chosen the individually appropriate FDs, the participants conducted the main pitch discrimination experiment, which is outlined and discussed in the following chapter.

4 Main Experiment: Discrimination of Temporal Pitch

This chapter forms the core of this thesis. The methods, i.e., the listener-specific setups, are presented, the data collected depicted and statistically analyzed. The outcome is discussed in order to subsequently draw general conclusions.

4.1 Methods

Procedure. Listeners, apparatus, stimuli, task, procedure, and fitting were identical to pitch discrimination pretests. The pulse properties of the LR stimuli were identical to the high-rate signals. Each of the six LR stimuli (three pairs with mean rates of 62.5, 125, and 250 pps, see Sec. 4) were loudness balanced to the 2000-pps high-rate reference using the adaptive procedure described in Section 3.2. For the pseudo-syllable stimuli, all combinations of MD (0.1, 0.3, 0.5, 0.7, and 0.9), SIPI (NS, FRS, and HRS) were measured (i) for both nominal F0s with the 2000-pps carrier and (ii) for the low nominal F0 with the 1000-pps carrier. The data for the different carrier rates was collected separately. Apart from CI77, the data for the 2000-pps carrier was collected together for both nominal F0s. For CI77, the two nominal F0s were tested separately due to time restrictions.

LR signals were tested with pulse rates of 62.5, 125, and 250 pps. They were repeated 200 times and included in the experimental blocks for the 2000-pps carrier. For 62.5 and 125 pps, the FD for the low nominal F0 was used. The 250 pps low-rate condition was tested with the FD for the high nominal F0.

Data Analysis. The data are analyzed using ANOVAs. Those require both normally distributed residuals and equal variances (homoscedasticity) across groups. Although ANOVAs are robust against violations of these requirements, significant ANOVA results that are based on requirement-violating data need to be double-checked with non-parametric methods to minimize type I errors. In order to test whether the data are suitable for ANOVA, we used Kolmogorov-Smirnov (K-S) tests with Lilliefors normal distribution significance correction certain groups. In order to assess the homogeneity of variances across different, we performed one-way ANOVAs on the absolute residuals with the same groups as factor levels (Levene's test).

For all statistical tests, the significance level was 5 %. Regularly, post-hoc tests following the omnibus ANOVAs were two-sided Tukey-HSD multiple comparison t-tests. They allow to compare every possible combination of factor levels. In case d's from high-rate conditions were compared to a LR condition, Dunnett's one-sided [14] (probe < control) multiple comparison t-tests were used, allowing to compare a control (LR condition) to more than one probe (high-rate conditions). The multiple comparison procedures adjust the pairwise significance levels such that the familywise significance level is 5 %.

4.2 Results

The collected data were evaluated in three stages: first, the 2000-pps data were analyzed, then the 2000-pps data for the low nominal F0 was compared with the 1000-pps data. Finally, the 2000-pps data were statistically compared with the LR.

2000-pps Carrier Rate

Prerequisites for ANOVAs. K-S tests for each of the six groups resulting from all possible combinations of nominal F0 and SIPI (NS, FRS, and HRS), pooled across listeners and MDs, revealed that the normality of residuals was only violated in the HRS condition for the high nominal F0 [$D(25) = 0.22$, $p < 0.01$]. Levene's test performed on the same six groups revealed a significant effect of the group [$F(5,144) = 5.32$, $p < 0.001$], indicating that at least one group was violating the assumption of homoscedasticity. Indeed, post-hoc Tukey-HSD tests [15] revealed a significantly higher variance for the same group that also violated the normality assumption, compared to all other groups. Thus, significant effects found in ANOVAs containing data from this group will also be investigated with non-parametric substitute procedures.

---

[14] We hypothesized that LR signals never perform worse than high-rate signals across the population. Visual inspection of the averaged results further provides no evidence that this assumption is violated in the pitch discrimination data. Together, we see the use of a one-sided test with probe < control as justified.

[15] Until further notice, all post-hoc tests mentioned are of this kind.

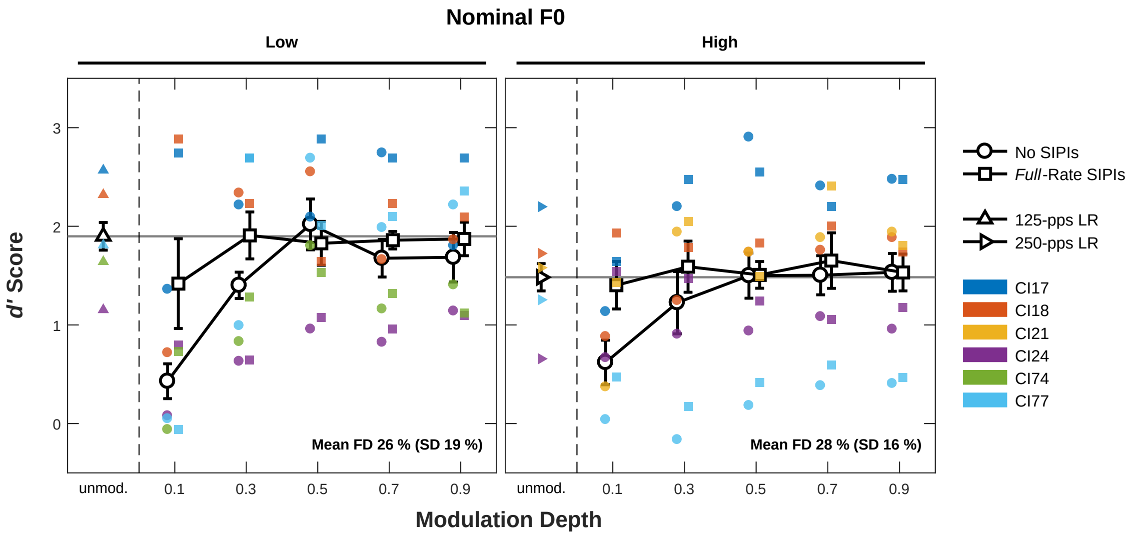

Figure 4.1 – Pitch discrimination d's for FRS and NS, 2000-pps carrier and both nominal F0s. LR conditions correspond to the F0s. Individual data are depicted in color. Error bars show normalized SEMs (Cousineau 2005, Morey 2008).

Full-Rate SIPIs. Figure 4.1 displays the pitch discrimination d's for FRS, NS, and the LR conditions as functions of MD. With NS, overall performance improved with increasing MD at both nominal F0s, starting at chance level and approaching the LR performance. With FRS, performance improved at low MDs for both nominal F0s compared to NS and it approached LR performance also at low MDs [16]. The results suggest no systematic deterioration of performance for FRS compared to NS. The overall high-rate performance never exceeded the LR performance.

In order to statistically investigate the effect of FRS, two-way repeated-measures (RM) ANOVAs were performed on the d's, separated for each nominal F0 and including the main factors MD (0.1, 0.3, 0.5, 0.7, and 0.9) and SIPI (NS and FRS) and their interaction. For the low nominal F0, both main factors were significant, MD [$F(4,36) = 6.57$, $p < 0.001$] and SIPI [$F(1,36) = 5.90$, $p < 0.05$], whereas their interaction was not [$F(4,36) = 2.05$, $p = 0.1074$]. Post-hoc tests on the main factors revealed significantly worse performance for MD 0.1 compared to all other MDs tested. The introduction of FRS significantly improved performance. Post-hoc tests on the non-significant interaction showed that NS performance was significantly worse at MD of 0.1 compared to MDs

---

[16] The benefit of FRS at MD 0.1 seems to be smaller at the low compared to the high nominal F0, but note also the high variance there.

≥ 0.5. With FRS, performance did not differ significantly across MDs. Further, the tests also did not show a significant benefit of FRS at any MD.

For the high nominal F0, both main factors were significant, MD [$F(4,36) = 6.04$, $p < 0.001$] and SIPI [$F(1,36) = 9.71$, $p < 0.01$], and their interaction was significant [$F(4,36) = 3.07$, $p < 0.05$]. Post-hoc tests revealed significantly worse NS performance at MD 0.1 compared to MDs ≥ 0.5. FRS performance did not differ significantly across MDs. At MD 0.1, there was a significant benefit due to FRS compared to NS.

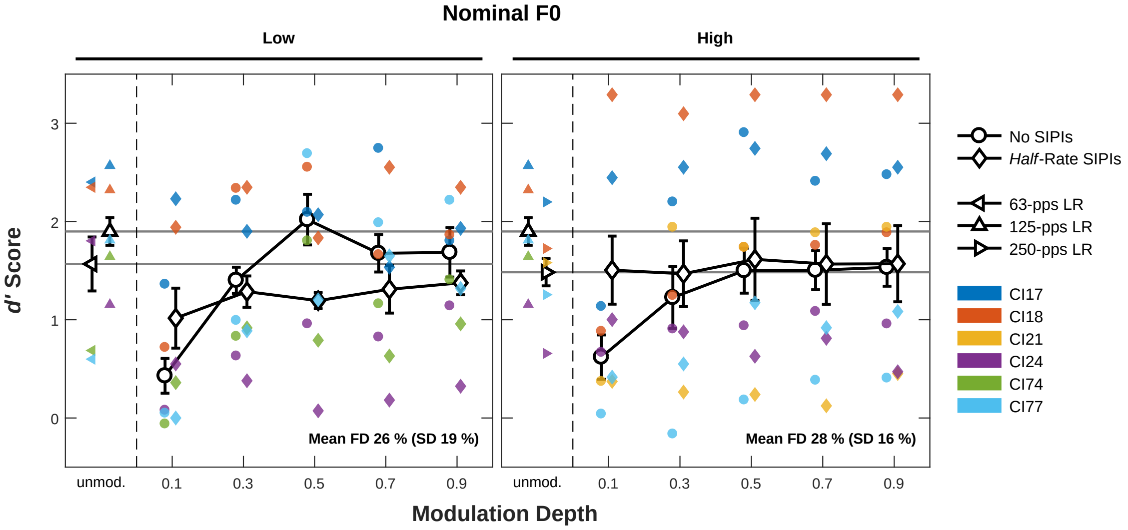

Figure 4.2 – Pitch discrimination d's for HRS and NS, 2000-pps carrier and both nominal F0s. The NS reference data are replicated from Fig. 4.1. LR conditions correspond to the F0 or the SIPI rate. Individual data in color. Error bars as in Fig. 4.1.

Half-Rate SIPIs. Figure 4.2 shows the pitch discrimination d's for HRS, NS, and the LR conditions as functions of MD. The reference data were only measured once and are thus identical to those above. Note that we compare LR conditions corresponding to the F0 and the SIPI rate. Visual inspection indicates substantial differences between the nominal F0s. Compared to NS, HRS performance at the low nominal F0 seems to be better at MD of 0.1 but worse for MDs ≥ 0.5. For the high nominal F0, d's were higher than for NS, however, the large error bars indicate a more heterogeneous performance across listeners as compared to all other high-rate conditions. For both nominal F0s, HRS performance does not seem to vary systematically across MDs.

Compared with LR, at the low nominal F0, HRS performance seems to be similar to the 63-pps LR condition representing the SIPI rate, but worse than the 125-pps LR condition representing the F0. At the high nominal F0, HRS performance seems to be more similar to the 250-pps LR condition representing the F0 than to the 125-pps LR condition representing the SIPI rate.

In order to quantify the effect of HRS, two-way RM-ANOVAs were conducted on the d's, separately for both nominal F0s, with main factors MD (0.1, 0.3, 0.5, 0.7, and 0.9) and SIPI(NS and HRS), and their interaction. For the low nominal F0, the main factor MD was significant [$F(4,36) = 7.18$, $p < 0.001$] whilst the main factor SIPI was not [$F(1,36) = 2.98$, $p = 0.0930$]. However, the interaction was significant [$F(4,36) = 3.71$, $p < 0.05$]. Post-hoc tests on the factor MD showed a significantly worse performance for MD of 0.1 compared to all other MDs. Post-hoc tests on the interaction showed that HRS never performed better than NS when comparing for the same MDs.

For the high nominal F0, none of the main effects was significant, MD [$F(4,36) = 1.12$, $p = 0.3634$] and SIPI [$F(1,36) = 2.21$, $p = 0.1459$]. The same also applied for the interaction [$F(4,36) = 0.77$, $p = 0.5533$]. Beyond that, none of the post-hoc comparisons on either the main effects or the interaction showed any significant outcomes.

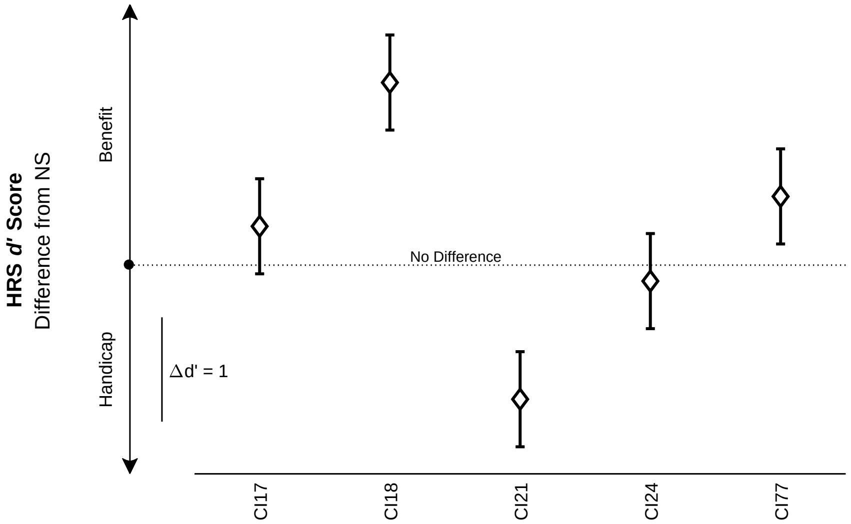

Figure 4.3 – Post-hoc Tukey-HSD tests following a one-way ANOVA with factor listener conducted on the d' differences HRS minus NS at the high nominal F0. Error bars represent confidence intervals such that the familywise significance level is 5 %.

To find out more about the between-subject differences at the high nominal F0, i.e., the individual benefit or handicap due to HRS compared to NS, a one-way ANOVA with factor listener (CI17, CI18, CI21, CI24, and CI77) was conducted on the difference between the HRS and the NS d' scores. The difference was calculated for each MD and then pooled for each listener across all MDs. The effect was significant [$F(4,20) = 26.68$, $p < 0.0001$]. Figure 4.3 shows the results of the post-hoc tests following the ANOVA. They revealed a significantly higher benefit for CI18 compared to all other listeners on the one hand, and a significantly higher handicap for CI21, compared to all other listeners, on the other hand. In between, the performance of the other three listeners did not significantly differ among each other.

The results in Figure 4.3 can also be analyzed from an absolute perspective to investigate whether individual listeners benefited from HRS. A significant benefit is indicated by confidence intervals above the “No Difference“ line. Thus, CI18 and CI77 perform significantly better with HRS than with NS, CI21 performs significantly worse. The performance of CI17 and CI24 does not change significantly with HRS.

Influence of nominal F0 on SIPI performance. In order to investigate potential effect of the nominal F0s, a two-way RM-ANOVA was conducted including the main factors nominal F0[17] (low/125 Hz and high/250 Hz) and SIPI (FRS and HRS) as well as their interaction. The main effect of nominal F0 was not significant [$F(1,91) = 0.02$, $p = 0.8858$], but the main effect of SIPI [$F(1,91) = 5.60$, $p < 0.05$] and the interaction reached significance [$F(1,91) = 5.93$, $p < 0.05$]. At the low nominal F0, post-hoc tests showed significantly worse performance with HRS. At the high nominal F0, however, performance did not differ significantly across SIPI conditions.

Because at the high nominal F0, the HRS data violated the ANOVA assumptions the main effect of SIPI was revisited using the Wilcoxon signed-rank test with two paired groups (FRS and HRS). In line with the ANOVA, SIPI yielded significant differences in performance ($Z = 2.17$, $p < 0.05$) with FRS being better than HRS. The interaction between SIPI and nominal F0 was investigated using a Friedman test for a one-way ANOVA

---

[17] In the pitch discrimination pretest, group statistics on the chosen FDs (Tab. 3.4) showed that the average FDs only marginally differ across nominal F0s and that this difference is overlaid by a large variance. We therefore regard the influence of FD on the results as negligible and the data as comparable across nominal F0s.

with RM on ranks and four groups resulting from all possible combinations of nominal F0 (low/125 Hz and high/250 Hz) and SIPI (FRS and HRS). The effect was significant [$\chi^2(3) = 9.58$, $p < 0.05$], again being in line with the ANOVA result. In summary, the non-parametric statistical tests fully support the results from the ANOVA, despite the violations of the prerequisites.

1000-pps Carrier Rate

Figure 4.4 displays the d' for the carrier rate of 1000 pps for FRS and Figure 4.5 for HRS, respectively. Visual inspection of the d's for the SIPI conditions yields a very similar picture compared to the 2000-pps carrier rate. In contrast to that, the NS performance seems to yield a better performance at 1000 pps than at 2000 pps, especially for low MDs, as shown by CI18 who, for example, showed maximum performance already at a MD of 0.1. The considerable variance in the MD 0.1/NS condition is also represented by the large error bar.

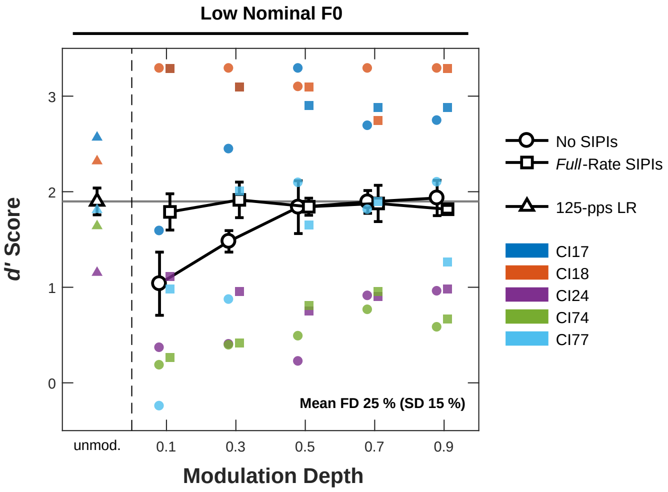

Figure 4.4 – Pitch discrimination d's for FRS and NS. 1000-pps carrier and low nominal F0. LR condition corresponds to the F0. Individual data in color. Error bars as in Fig. 4.1.

In order to examine the differences in the data across both carrier rates and thus a potential influence of the carrier on the pitch discrimination performance, a two-way RM-ANOVA with main factors carrier rate (1000 pps and 2000 pps) and SIPI (NS, FRS, and HRS), and their interaction was conducted. The main factor carrier rate was not significant [$F(1,140) = 0.53$, $p = 0.4662$] but the main factor SIPI was significant [$F(2,140) = 15.71$, $p < 0.0001$]. The interaction was not significant [$F(2,140) = 0.73$, $p = 0.4822$]. Post-hoc tests show both significantly better performance with FRS and significantly worse performance with HRS compared to NS.

Comparison to Low-Rate Conditions

By comparing each of the three high-rate signal types (NS, FRS, and HRS) to the LRs, we aimed to quantify the SIPI performance more than by only investigating relative differences. For these comparisons, only the 2000-pps data are used, separately for each nominal F0.

Low nominal F0. We ran a one-way RM-ANOVA with factor MD on the d's, including NS data for all five MDs and the 125-pps LR. In the technical implementation of the test, the latter was encoded as MD of 0. The effect was significant [$F(5,20) = 11.25$, $p < 0.001$]. Dunnett's post-hoc tests[18] were used to compare the high-rate conditions to the LR acting as a reference. We found significantly worse performance for MD 0.1/NS. All other MDs did not differ significantly from the LR.

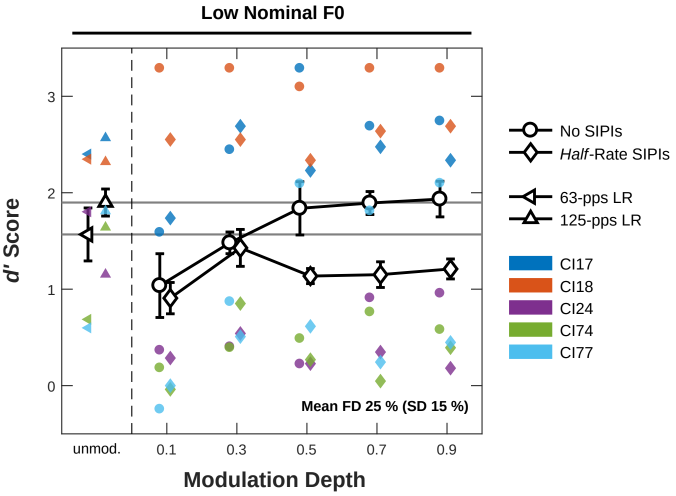

Figure 4.5 – Pitch discrimination d's for HRS and NS. 1000-pps carrier and low nominal F0. The NS data are replicated from Fig. 4.4. LR conditions correspond to F0 or SIPI rate. Individual data in color. Error bars as in Fig. 4.1.

In order to search for differences between various conditions at MD of 0.1, we conducted a one-way RM-ANOVA including the factor condition with the following levels: 125-pps LR, MD 0.1/NS, MD 0.1/FRS, and MD 0.1/HRS. The effect of the condition was significant [$F(3,12) = 9.29$, $p < 0.01$]. Post-hoc tests comparing all MDs to the LR condition again showed significantly worse performance with NS compared to the LR. Further, they showed no significant difference between FRS and LR, but significantly worse performance for HRS than for LR.

---

[18] Again, until further notice, all following post-hoc tests are set up in the same way.

Since HRS may encode two frequency cues (the F0 of 125 Hz in the AM and the SIPI rate of 63 pps), a comparison to a 125-pps LR might have been misleading. Thus, we repeated the first ANOVA with factor MD, but this time with the HRS d's for all five MDs and the 63-pps instead of the 125-pps LR condition, again encoded as MD of 0. The effect was not significant [$F(5,20) = 0.85$, $p = 0.5300$]. Post-hoc tests comparing all MDs to the LR condition showed no significant differences between HRS and LR for any MD.

High nominal F0. First, we compared the NS d's for all MDs with the 250-pps LR. Hence, we conducted a one-way RM-ANOVA with factor MD. The LR was encoded as MD of 0. The effect of MD was significant [$F(5,20) = 5.07$, $p < 0.01$]. Post-hoc tests comparing all MDs to the LR revealed a significantly worse performance for MD of 0.1.

To search for differences across the SIPI conditions at MD of 0.1, we ran a one-way RM-ANOVA with factor condition including the following levels: 250-pps LR, MD 0.1/NS, MD 0.1/FRS, and MD 0.1/HRS. The effect was not significant [$F(3,12) = 2.76$, $p = 0.0880$]. Post-hoc tests comparing the MDs to the LR condition showed significantly worse performance for NS but not for FRS and HRS.

To compare the HRS d' to the 125-pps LR condition representing the SIPI rate, we repeated the first ANOVA with factor MD, but with the HRS d's for all MDs and the LR encoded as MD of 0. The effect was not significant [$F(5,19) = 0.17$, $p = 0.9710$]. Post-hoc tests comparing the MDs to the LR condition showed no significant differences between HRS and LR.

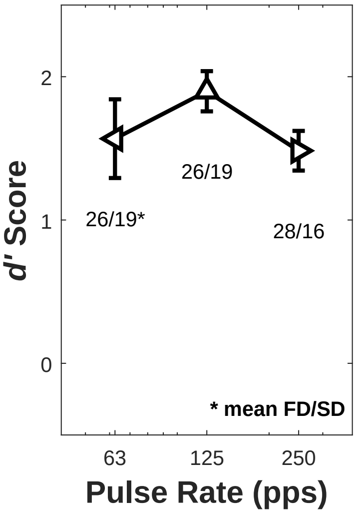

Figure 4.6 – Pitch discrimination d's with the LR conditions averaged across listeners. Error bars as in Fig. 4.1.

Low-rate signals. Figure 4.6 shows the $d's$ as a function of pulse rate. To examine the influence of the pulse rate on pitch discrimination with LR, we ran a one-way RM-ANOVA with factor pulse rate (63, 125, and 250 pps). The effect was not significant [$F(2,7) = 1.79$, $p = 0.2340$].

4.3 Discussion

After providing a substantial amount of data and statistical analysis, the results are now discussed in the light of related literature.

The Influence of MD and SIPI on Performance with Pseudo-Syllable Signals

No SIPI. In our experiment, NS d's monotonically increased with increasing MD for both nominal F0s, starting at chance level for MD 0.1 and converging at a supra-threshold score ($d' \geq 1$) for MD 0.9. This seems reasonable when assuming that increasing the MD increases the salience of AM-coded pitch.

McKay et al. (1994, 1995) have shown that the pitch associated with AM stimuli can be seen as a weighted average of carrier and modulation rate with MD being the weighting factor. Thus, for high carrier rates sufficiently above the rate limit for temporal pitch, perception for low MDs can reasonably be assumed to be weak. With increasing MD, the perceived pitch converges towards the modulation rate and, in case the latter is sufficiently below the rate limit, pitch discrimination performance should also increase. The model by McKay and colleagues is thus capable of explaining the NS performance in our experiment. Geurts and Wouters (2001) measured single-channel pitch discrimination performance with SAM stimuli and MDs ranging from 0.05 to 0.99. They concluded that optimal performance was achieved for MDs $\geq$ 0.2 and that pitch coding is weak for lower MDs. This is in line with our findings showing that pitch coding is at chance level for MD of 0.1.

In everyday situations, noise in the environment reduces the MD of CIS-like signals. Further, currently available CI processors produce signals with MDs ranging only between approximately 0.1 and 0.4 (cf. Sec. 2.3). Thus, AM-coded pitch is expected to be be poor in everyday situations.

Full-Rate SIPI. Compared to NS, FRS performance was significantly better across F0s. For the high nominal F0s, it was also significantly better in a separate analysis at MD of 0.1 only. Further, performance did not differ significantly across MDs and was thus homogeneous and robust across different conditions. When AM-based pitch is weakened due to noise, FRS allows to convey F0 and other pitch cues as long as the F0-modulation can be detected by the signal processing algorithms. Focusing on F0 and its typical range from roughly 80 to 300 Hz for male and female speakers [19], FRS seems to be is capable of providing F0 cues with gender-independent salience.

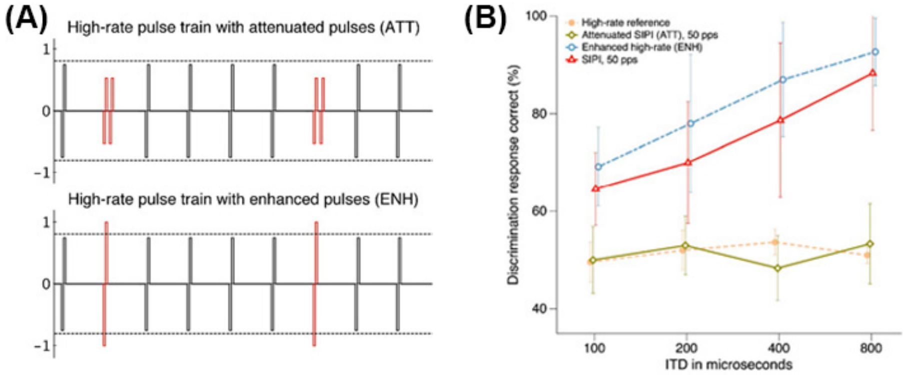

Figure 4.7 – Assessing the meachanism underlying the SIPI effect. **(A)** Comparison signals and **(B)** results. SIPI width 6 %, SIPI rate 50 pps, error bars show 95 % confidence intervals (Srinivasan et al. 2018).

To reveal the mechanism underlying the SIPI effect, Srinivasan et al. (2018) investigated the ITD discrimination performance of an unmodulated 1000-pps carrier rate, 50-pps SIPI rate, and 6-% SIPI fraction pulse train. They compared the SIPI condition to two further conditions: (1) attenuated SIPI pulse trains (ATT, the short-term power of an ATT pulse pair is equal to one regular carrier pulse) and (2) enhanced high-rate pulse trains (ENH, the short-term power of one ENH pulse is equal to a regular SIPI pulse pair) [Fig. 4.7 (A)]. Results [Fig. 4.7 (B)] showed comparable performance for regular SIPI pulse trains and ENH pulse trains suggesting that both these stimuli introduce an artificial internal AM. Because the ATT performance did not show improved sensitivity compared to the high-rate reference without SIPI, they could not confirm a hypothesis that SIPIs enhance performance by some mechanism depending on their particular temporal pattern.

As Srinivasan et al. (2018) could not attribute the SIPI effect to a mechanism depending on the temporal pattern, it is also reasonable to assume the same for pitch discrimination. Thus, the model of the artificial internal AM is also capable of explaining the improved temporal pitch sensitivity and the constant performance across MDs.

---

[19] Proposal of National Institutes of Health (NIH) project “Bilateral Cochlear Implants: Physiology and Psychophysics“, grant #: R01DC005775.

Assuming that SIPIs enhance the internal AM, it seems reasonable to also compare the FRS to signal processing strategies that modify the AM already externally, e.g., by modifying the envelope shape, as done by Wouters and colleagues (Geurts and Wouters 2001, Laneau et al. 2006, Milczynski et al. 2009). They found good single-channel performance, but also detrimental effects on speech understanding in multi-channel configurations. The latter finding is highly plausible as modifications of the envelope in favor of F0 are likely to distort the spectral information encoding speech. By contrast, the SIPI approach has little impact on the signal envelope and might thus be more suitable to convey both speech and pitch cues. Behavioral data investigating this hypothesis is not available to date and experiments testing speech perception with SIPIs may shed light. However, such experiments need to consider multi-channel processing, as regular speech processing requires more than one channel. Previous studies investigating AM frequency discrimination suggested a beneficial effect of multi-channel AM pitch compared to single-channel AM pitch(see, e.g., Galvin et al. 2015), thus, in such experiments, data of multi-channel SIPI pitch would also be of interest.

Half-Rate SIPI. HRS did not differ significantly from NS, both across MDs and nominal F0s. As it seems, including a second lower pitch cue did not provide a robust benefit which would have manifested in a homogeneous performance. However, it also never showed significant detrimental effects. In HRS conditions which are beneficial for, e.g., ITD perception, it would thus likely not degrade pitch perception and not introduce a trade-off between pitch and other percepts. At the high nominal F0, the variance was significantly higher compared to all other high-rate signal conditions. This occurred due to considerable inter-individual differences in benefit, with CI18 showing exceptionally high sensitivity and CI21 showing remarkably low sensitivity.

In situations with two potential pitch cues, the question which of them is dominant seems natural. Carlyon and colleagues (Carlyon et al. 2002, van Wieringen et al. 2003) performed single-channel experiments with both unmodulated and AM “irregular-rate“ stimuli (Fig. 4.8). For unmodulated signals, Carlyon et al. (2002) found a dominance of (i) the first-order intervals[20], (ii) the frequency corresponding to the “common“ interval (Carlyon 1997), and (iii) a dominance of the interval corresponding to the lower

---

[20] A first-order interval is the interval between two subsequent pulses. In Carlyon et al. (2002), that is either 4 or 6 ms.

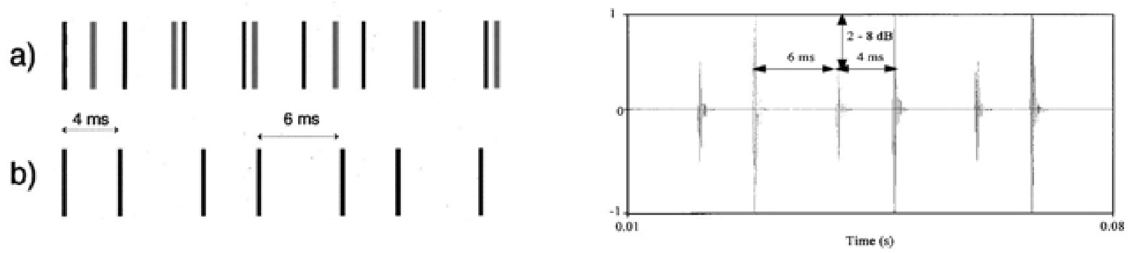

Figure 4.8 – Irregular-rate “4-6“ stimuli. *Left:* Unmodulated pulse trains with (a) one common among numerous other intervals and (b) two equally occurring intervals (Carlyon et al. 2002). *Right:* Acoustic version of an AM stimulus where the MD is expressed as the difference between the peaks of the two pulses in dB (van Wieringen et al. 2003).

frequency when two intervals are occurring equally often. For AM signals, van Wieringen et al. (2003) reported that in case of a high MD the pitch was close to the second-order interval[21] whereas it converged towards the pitch corresponding to the lower first-order interval with decreasing MD (consistent with the data for unmodulated signals), suggesting some kind of “mean rate“ cute[22] between these the second-order interval and the lower first-order interval (Carlyon 1997).

In the light of these findings, our SIPI stimuli were regular signals with all intervals occurring equally often. Our SIPI rate corresponds to the second-order interval and the F0 corresponds to (both) first-order intervals. The unmodulated signals from Carlyon et al. (2002) suggest a tendency towards lower first-order interval (6 ms) pitches, the AM signals suggest an increase of the pitch from that corresponding to the second-order interval (10 ms) towards that corresponding to the lower first-order interval (6 ms) with a decrease of the MD. Note that our definition of the MD (relative to the DR) is different than that from van Wieringen et al. (2003). Thus, our NS data showed increased salience of the first-order interval corresponding to the F0 with increasing MD, whereas data from van Wieringen et al. (2003) showed decreased salience of the first-order interval with increasing MD.

In our experiment, with a MD of 0.1, the F0 cue in the envelope was expected to be very weak. Hence, the HRS pitch might rather correspond to the second-order interval (SIPI rate^{-1}). Following the AM irregular-rate results from van Wieringen et al. (2003), with

---

[21] A second-order interval is the interval between a pulse and the next but one pulse. In van Wieringen et al. (2003), that corresponded to an interval of 10 ms.

[22] Within a certain period of time, the total number of pulses is proportional to the mean rate.

increasing MD, this pitch should increase towards the pitch of the F0 as a consequence of averaging the second-order and the first-order interval ($F0^{-1}$).

Based on our results that HRS performance is similar across MDs, it seems reasonable to assume that the pitch does not change when varying the MD. Considering that HRS contains two frequency cues, this would result in the listeners either focusing on the F0 or on the SIPI rate constantly across MDs.

The hypothesis of a dominant SIPI rate cue can be used to explain the HRS results for the high nominal F0. If the SIPI rate (being half of the F0) dominates perception, the pitch will be effectively lowered by one octave, shifting the pitch discrimination sensitivity to the low nominal F0. For example, CI18, who showed an exceptional benefit (cf. Fig. 4.3) of HRS compared to NS, effectively discriminated F0s with a FD of 13 % at a nominal F0 of 125 Hz. This is twice the FD chosen for him for the low nominal F0 based on his pitch discrimination pretest results. The discrimination would likely be rather easy and it might thus be that CI18 focused on the SIPI rate. On the other side, CI21 was not sensitive at all the low nominal F0 and was thus excluded for this nominal F0 after the pitch discrimination pretest. If CI21 was focusing on the SIPI rate, results would have been based on an F0 of 125 Hz instead of 250 Hz. Hence, this could explain the huge handicap introduced by HRS (absolute performance is at chance level). For the other three listeners, the overall sensitivity was generally worse than CI18's and the relative differences of the FDs chosen in the pitch discrimination pretest were smaller. Thus, pitch shifts from the high to the low nominal F0 probably had a smaller impact on their discrimination performance.

Differences across SIPI Approaches. Compared to FRS, the results for HRS pitch discrimination are much more ambiguous. Whilst FRS is a straightforward way to encode pitch and is thus also motivated from a pitch-based perspective, HRS was motivated from a different point of view, namely ITD perception and rate limitation. At the high nominal F0, FRS and HRS were still similar. Contrary to that, HRS was significantly worse at the low nominal F0. Based on that, FRS seems to be generally preferable over HRS, in case there are no other restrictions. However, if restrictions have to be made, HRS can be used for higher F0s. In such a case, considerable individual differences in the benefit are expected, likely mediated by the individual sensitivity in the region of the SIPI rate.

Comparison of Pseudo-Syllable and Low-Rate Signals

When comparing NS with LR data, we found no significant differences for MDs ≥ 0.3. When comparing FRS with LR data, we found no significant differences for any MDs. When comparing HRS with LR data, at the low nominal F0, listeners performed significantly worse with HRS than with low rates corresponding to the F0. HRS performance did not differ significantly from the low rate corresponding to the SIPI rate. At the high nominal F0, we found no significant differences between HRS and low rates corresponding to both F0 and SIPI rate.

Taken together, our results suggest that the SIPI rate dominated pitch perception.

Upper Limit of Temporal Pitch

The FRS and HRS conditions were measured at two nominal F0s. The main effect of F0 was not significant indicating no evidence for a rate limitation. Further, at the high nominal F0, an F0 potentially affected by a rate limitation, the comparison between FRS and HRS did not show any significant differences. Thus, we cannot tell whether HRS is capable of circumventing a rate limitation at all.

We further measured discrimination performance with LR signals to assess whether our data might be affected by some form of rate limitation. Also this analysis did not show any effect of pulse rate on LR performance, further suggesting no rate limitation in our study.

Ihlefeld et al. (2015) compared the rate limitation between rate-pitch and ITD discrimination by using standard pulse rates of 100 to 500 pps and a fixed FD of 35 %[23] at three tonotopic locations. They found a significant effect of rate. As shown in Figure 4.9, their rate-pitch performance (red and blue curves) began to decrease for rates between 200 and 300 pps[24]. It is thus likely that our F0s did not severely touch our participants' individual rate limits.

---

[23] Note that the FD is defined *re* the standard rate.

[24] ITD performance (black curve) already began to decrease for rates between 100 and 200 pps.

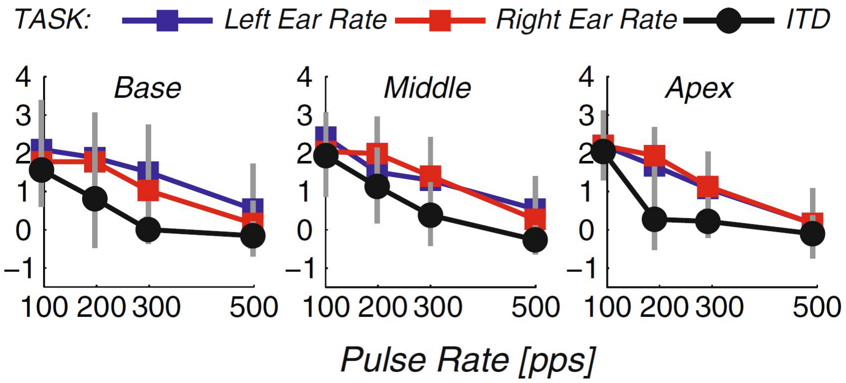

Figure 4.9 – Monaural rate-pitch and ITD d's as a function of pulse rate for three cochlear locations (Ihlefeld et al. 2015).

The Influence of the Carrier Rate

Because our statistical analysis did not show a significant effect of carrier rate, we cannot find a significant influence of the carrier rate on AM pitch discrimination performance. However, we measured 1000-pps performance at low F0s only. Thus, an effect of the carrier at other F0s cannot be ruled out. As current CIS-like stimulation strategies typically use carrier rates within the range we tested, our findings should be relevant for clinical applications.

Green et al. (2012) measured JNDs for modulation rate detection as a function of pulse rates with rates ranging from 482 to 5787 pps. They could not find a significant influence of pulse rate on the JNDs, which is in line with us not finding significant differences in discrimination performance.

5 Conclusions and Outlook

Within this thesis, we measured pitch discrimination performance with pseudo-syllable signals encoded by two differently motivated SIPI approaches (FRS and HRS). Two parameters were systematically varied: MD, ranging from noisy everyday situation to artificial laboratory situation, and nominal F0 representing male and female voices. We focused on a 2000-pps carrier rate to sufficiently oversample all F0s, but also collected data for a 1000-pps carrier rate in order to compare to CIS-like stimulation and previous SIPI-ITD studies, both having a similar lower carrier rate. As references, we also measured performance for pseudo-syllable signals without SIPI pulses and unmodulated pulse trains with low rates corresponding to the SIPI rates used. The findings can be summarized as follows:

(1) High-rate stimulation without SIPIs deteriorates with decreasing MD which mainly represents everyday situations where the effective MD is lowered due to environmental noise. Compared to the low rates, high-rate performance is significantly worse at low MDs.

(2) The insertion of SIPI pulses at full rate seems to be a straightforward way to encode the F0 cue regardless of MD, yielding significant improvements at low MDs and no detrimental effects at high MDs. This conclusion seems to hold for both male and female voices as supported by the results for two (nominal) F0 ranges. Compared to the low rates, we did not find significant performance differences at any MD and F0. This suggests that SIPI insertion at full rate restores pitch discrimination performance with pseudo-syllable signals particularly at low MDs.

(3) If the maximum SIPI rate needs to be limited, the insertion of SIPI pulses at half of the F0 might still provide benefit. Although the F0 cue might be smeared being represented by two ambiguous frequencies in the signal, the HRS approach might still provide access to F0 variations needed, e.g., for prosody perception. However, the benefit was expected to be less clear. At low F0s, HRS showed rather detrimental effects, at least for moderate to high MDs. At high F0s, the benefit was highly listener-specific with one listener benefiting a lot and one listener showing much worse performance compared to the NS data. Compared to the low rates, our results suggest to favor the SIPI rate as the dominant cue. Im summary, the results might reflect a SIPI-induced

pitch shift. Unfortunately, our experiment was not designed to determine the absolute pitch and more investigations are necessary.

(4) For both SIPI signals and low rates, we did not find any significant effect of F0 or pulse rate, respectively, on pitch discrimination performance. There is therefore no evidence for a rate limitation in our experiment. Consequently, we cannot determine if the HRS approach did circumvent a rate limitation.

(5) Pitch discrimination at the low nominal F0 did not significantly vary with the carrier rate, at least between 1000 and 2000 pps.

In general, we have shown that the insertion of SIPI pulses can have a beneficial effect on temporal pitch perception in electric hearing. In order to maximize the advantages and minimize the drawbacks, our results suggest to insert SIPI pulses at full rate for both male and female F0s. If this is not acceptable, e.g., due to detrimental effects of higher SIPI rates on other percepts, we suggest to limit the SIPI rate to, e.g., half of the F0. However, detrimental effects are likely to occur for very low SIPI rates, e.g., when constantly inserting SIPIs with half rate at lower F0s.

Implications for Signal Coding Strategies. Srinivasan et al. (2018) have shown that SIPI insertion improves ITD sensitivity with unmodulated high-rate signals. Preliminary data (Srinivasan et al. 2017) also suggest a beneficial effect with pseudo-syllable (AM) signals. Combining these results with our results, a joint improvement of both ITD and pitch perception while maintaining high speech intelligibility seems likely to be feasible. However, the actual influence of SIPI insertion on speech understanding with CIS-like stimulation remains to be shown in future studies.

References

ANSI (1994), American National Standard Acoustical Terminology, ANSI S1.1-1994, Standard, American National Standards Institute, New York.

Arnoldner, C., Riss, D., Kaider, A., Mair, A., Wagenblast, J., Baumgartner, W.-D., Gstöttner, W. and Hamzavi, J.-S. (2008), 'The Intensity-Pitch Relation Revisited: Monopolar Versus Bipolar Cochlear Stimulation', *Laryngoscope* **118**(9), 1630–1636.

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 Hear* **25**, 275–283.

Bendor, D., Osmanski, M. S. and Wang, X. (2012), 'Dual-Pitch Processing Mechanisms in Primate Auditory Cortex', *J. Neurosci.* **32**(46), 16149–16161.

Bendor, D. and Wang, X. (2005), 'The neuronal representation of pitch in primate auditory cortex', *Nature* **436**(7054), 1161–1165.

Boulet, J., White, M. and Bruce, I. C. (2015), 'Temporal Considerations for Stimulating Spiral Ganglion Neurons with Cochlear Implants', *J. Assoc. Res. Otolaryngol.* **17**(1), 1–17.

Brochier, T., McKay, C. and McDermott, H. (2018), 'Rate modulation detection thresholds for cochlear implant users', *J. Acoust. Soc. Am.* **143**(2), 1214–1222.

Carlyon, R. P. (1997), 'The effects of two temporal cues on pitch judgments', *J. Acoust. Soc. Am.* **102**(2), 1097–1105.

Carlyon, R. P., Lynch, C. and Deeks, J. M. (2010), 'Effect of stimulus level and place of stimulation on temporal pitch perception by cochlear implant users', *J. Acoust. Soc. Am.* **127**(5), 2997–3008.

Carlyon, R. P. and Shackleton, T. M. (1994), 'Comparing the fundamental frequencies of resolved and unresolved harmonics: Evidence for two pitch mechanisms?', *J. Acoust. Soc. Am.* **95**(6), 3541–3554.

Carlyon, R. P., van Wieringen, A., Long, C. J., Deeks, J. M. and Wouters, J. (2002), 'Temporal pitch mechanisms in acoustic and electric hearing', *J. Acoust. Soc. Am.*

112(2), 621–633.

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

Cousineau, D. (2005), ‘Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson’s method’, *Tutor Quant Methods Psychol* **1**(1), 42–45.

Dorman, M. and Wilson, B. (2004), ‘The Design and Function of Cochlear Implants’, *Am. Sci.* **92**(5), 436–445.

Dynes, S. B. C. and Delgutte, B. (1992), ‘Phase-locking of auditory-nerve discharges to sinusoidal electric stimulation of the cochlea’, *Hearing Research* **58**(1), 79–90.

Friesen, L. M., Shannon, R. V., Baskent, D. and Wang, X. (2001), ‘Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants’, *J. Acoust. Soc. Am.* **110**(2), 1150–1163.

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.

Geurts, L. and Wouters, J. (2001), ‘Coding of the fundamental frequency in continuous interleaved sampling processors for cochlear implants’, *J. Acoust. Soc. Am.* **109**(2), 713–726.

Geurts, L. and Wouters, J. (2004), ‘Better place-coding of the fundamental frequency in cochlear implants’, *J. Acoust. Soc. Am.* **115**(2), 844–852.

Green, T., Faulkner, A. and Rosen, S. (2012), ‘Variations in Carrier Pulse Rate and the Perception of Amplitude Modulation in Cochlear Implant Users’, *Ear Hear.* **33**(2), 221–230.

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.

Hancock, K. E., Chung, Y., McKinney, M. F. and Delgutte, B. (2017), ‘Temporal Envelope Coding by Inferior Colliculus Neurons with Cochlear Implant Stimulation’, *J. Assoc. Res. Otolaryngol.* **18**(6), 771–788.

Hilbert, D. (1912), *Grundzüge einer allgemeinen Theorie der linearen Integralgleichungen*, Teubner.

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.

Jesteadt, W. (1980), 'An adaptive procedure for subjective judgments.', *Percept. Psychophys.* **28**, 85–88.

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

Kreft, H. A., Oxenham, A. J. and Nelson, D. A. (2010), 'Modulation rate discrimination using half-wave rectified and sinusoidally amplitude modulated stimuli in cochlear-implant users', *J. Acoust. Soc. Am.* **127**(2), 656–659.

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', *Proc. Natl. Acad. Sci. U.S.A.* **105**(2), 814–817.

Laback, B., Majdak, P. and Baumgartner, W.-D. (2007), 'Lateralization discrimination of interaural time delays in four-pulse sequences in electric and acoustic hearing', *J. Acoust. Soc. Am.* **121**(4), 2182–2191.

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

Laneau, J., Wouters, J. and Moonen, M. (2006), 'Improved Music Perception with Explicit Pitch Coding in Cochlear Implants', *Audiol Neurotol* **11**(1), 38–52.

Licklider, J. C. R. (1951), 'A Duplex Theory of Pitch Perception', *Experientia* **VII**(4), 128–134.

Lindenbeck, M. (2017), Pitch SIPI. Experiments for Rate Pitch Perception in Cochlear Implant Listeners with Short-Interpulse-Interval (SIPI) Stimulation, Audio engineering project thesis, Institute for Electronic Music and Acoustics, University of Music and Performing Arts Graz.
URL: *http://phaidra.kug.ac.at/o:66457*

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.

Macherey, O. and Carlyon, R. P. (2012), 'Place-pitch manipulations with cochlear implants', *J. Acoust. Soc. Am.* **131**(3), 2225–2236.

McDermott, H. J. (2004), 'Music perception with cochlear implants: a review', *Trends Amplif* **8**(2), 49–82.

McDermott, H. J. and McKay, C. M. (1997), 'Musical pitch perception with electrical stimulation of the cochlea', *J. Acoust. Soc. Am.* **101**(3), 1622–1631.

McKay, C. M. and Carlyon, R. P. (1999), 'Dual temporal pitch percepts from acoustic and electric amplitude-modulated pulse trains', *J. Acoust. Soc. Am.* **105**(1), 347–357.

McKay, C. M. and Henshall, K. R. (2009), 'Amplitude Modulation and Loudness in Cochlear Implantees', *J. Assoc. Res. Otolaryngol.* **11**(1), 101–111.

McKay, C. M., McDermott, H. J. and Clark, G. M. (1994), 'Pitch percepts associated with amplitude-modulated current pulse trains in cochlear implantees', *J. Acoust. Soc. Am.* **96**(5), 2664–2673.

McKay, C. M., McDermott, H. J. and Clark, G. M. (1995), 'Pitch matching of amplitude-modulated current pulse trains by cochlear implantees: The effect of modulation depth', *J. Acoust. Soc. Am.* **97**(3), 1777–1785.

Mehta, A. H. and Oxenham, A. J. (2017), 'Vocoder Simulations Explain Complex Pitch Perception Limitations Experienced by Cochlear Implant Users', *J. Assoc. Res. Otolaryngol.* **18**(6), 789–802.

Milczynski, M., Wouters, J. and van Wieringen, A. (2009), 'Improved fundamental frequency coding in cochlear implant signal processing', *J. Acoust. Soc. Am.* **125**(4), 2260–2271.

Morey, R. D. (2008), ‘Confidence Intervals from Normalized Data: A correction to Cousineau (2005)’, *Tutor Quant Methods Psychol* **4**(2), 61–64.

Oxenham, A. J. (2008), ‘Pitch Perception and Auditory Stream Segregation: Implications for Hearing Loss and Cochear Implants’, *Trends Amplif* **12**(4), 316–331.

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

Oxenham, A. J. (2013), ‘Revisiting place and temporal theories of pitch’, *Acoust. Sci. Technol.* **34**(6), 388–396.

Oxenham, A. J., Bernstein, J. G. W. and Penagos, H. (2004), ‘Correct tonotopic representation is necessary for complex pitch perception’, *P. Natl. Acad. Sci. U.S.A* **101**(5), 1421–1425.

Pijl, S. (1997), ‘Pulse rate matching by cochlear implant patients: effects of loudness randomization and electrode position.’, *Ear Hear.* **18**, 316–325.

Rose, J. E., Brugge, J. F., Anderson, D. J. and Hind, J. E. (1968), *Hearing Mechanisms in Vertebrates*, Churchill, London, chapter Patterns of Activity in Single Auditory Nerve Fibres of the Squirrel Monkey, pp. 144–157.

Schouten, J. F. (1940), ‘The residue and the mechanism of hearing’, *P. K. Ned. Akad. Wetensc.* **43**, 991–999.

Shackleton, T. M. and Carlyon, R. P. (1994), ‘The role of resolved and unresolved harmonics in pitch perception and frequency modulation discrimination’, *J. Acoust. Soc. Am.* **95**(6), 3529–3540.

Shannon, R. V. (1983), ‘Multichannel electrical stimulation of the auditory nerve in man. I. Basic psychophysics.’, *Hear. Res.* **11**, 157–189.

Shield, B. (2006), Evaluation of the social and economic costs of hearing impairment, resreport, Hear-It.

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’, *J. Acoust. Soc. Am.* **141**, 3973–3974.

Srinivasan, S., Laback, B., Majdak, P. and Delgutte, B. (2018), 'Introducing Short Interpulse Intervals in High-Rate Pulse Trains Enhances Binaural Timing Sensitivity in Electric Hearing', *J. Assoc. Res. Otolaryngol.* .

Steinschneider, M., Reser, D. H., Fishman, Y. I., Schroeder, C. E. and Arezzo, J. C. (1998), 'Click train encoding in primary auditory cortex of the awake monkey: Evidence for two mechanisms subserving pitch perception', *J. Acoust. Soc. Am.* **104**(5), 2935–2955.

Townshend, B., Cotter, N., Van Compernolle, D. and White, R. L. (1987), 'Pitch perception by cochlear implant subjects', *J. Acoust. Soc. Am.* **82**, 106–115.

van Hoesel, R. J. M. (2008), 'Observer weighting of level and timing cues in bilateral cochlear implant users', *J. Acoust. Soc. Am.* **124**(6), 3861–3872.

van Hoesel, R. J. M., Jones, G. L. and Litovsky, R. Y. (2009), 'Interaural Time-Delay Sensitivity in Bilateral Cochlear Implant Users: Effects of Pulse Rate, Modulation Rate, and Place of Stimulation', *J. Assoc. Res. Otolaryngol.* **10**(4), 557–567.

van Wieringen, A., Carlyon, R. P., Long, C. J. and Wouters, J. (2003), 'Pitch of amplitude-modulated irregular-rate stimuli in acoustic and electric hearing', *J. Acoust. Soc. Am.* **114**(3), 1516–1528.

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.

WHO (2018), 'Who calls for action from governments and their partners to stem the rise in hearing loss'. accessed Mar 15, 2018.
URL: *http://www.who.int/deafness/world-hearing-day/2018-note-to-media/en/*

Wilson, B. S. and Dorman, M. F. (2008), 'Cochlear implants: A remarkable past and a brilliant future', *Hearing Res.* **242**(1-2), 3–21.

Wilson, B. S., Finley, C. C., Lawson, D. T., Wolford, R. D., Eddington, D. K. and Rabinowitz, W. M. (1991), 'Better speech recognition with cochlear implants', *Nature* **352**(6332), 236–238.

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

Wouters, J., McDermott, H. J. and Francart, T. (2015), ‘Sound Coding in Cochlear Implants: From electric pulses to hearing’, *IEEE Signal Process. Mag.* **32**(2), 67–80.

Zeng, F.-G. (2002), ‘Temporal pitch in electric hearing’, *Hear. Res.* **174**(1-2), 101–106.

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

Glossary

AM	Amplitude modulation: Audio information is commonly transmitted by the CIs via AM.
ANOVA	Analysis of variance: Statistical method to compare the means of three or more groups using parametric data.
AN	Auditory nerve
BEHL	Better Ear Hearing Level (dB)
BM	Basilar membrane: structural element of the cochlea that separates the scala media and the scala tympani.
CIS	Continuous interleaved sampling: CI signal processing strategy being very successful in restoring speech perception in quiet (Wilson et al. 1991).
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 in μA or dB.
ENV	Slowly varying envelope of a time domain signal. It can be extracted using, e.g., the Hilbert transform (Hilbert 1912).
F0	Fundamental frequency: The frequency of a pure tone or a harmonic complex that is perceived as the tone’s pitch.

FD	Frequency difference: The difference between two two frequencies expressed in percent, i.e., $\mathrm{FD} = 100 \cdot \left(\frac{f_{\mathrm{higher}}}{f_{\mathrm{lower}}} - 1 \right) \qquad (\%).$
FM	Frequency modulation
FRS	SIPI insertion with full rate: One extra pulse is inserted every AM period. F0 and SIPI rate are equal.
FS	Rapidly varying fine structure of a time domain signal.
GM	Geometric mean
HI	Hearing impairment
HRS	SIPI insertion with half rate: One extra pulse is inserted every other AM period of a pseudo-syllable signal. The SIPI rate is half of the F0.
IPI	Interpulse interval: The inverse of the carrier pulse rate.
ITD	Interaural time difference: ITDs are the dominant cue for localizing sounds containing low frequencies.
JND	Just Noticeable Difference: This quantity describes the amount by which a parameter has to be changed in order to cause a variation in perception.
LR	Low-rate pulse trains coding pitch via the pulse repetition rate.
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).
NH	Normal-hearing: NH listeners hear without impairment.
NS	Pseudo-syllable signal without SIPIs, acting as a high-rate reference.
pps	pulses per second: pps is a unit for pulse repetition rates.

RM	Repeated measures: Each participant is tested on the same setup (within-subjects design), normally with a high amount of repetitions per setup condition. Thus, the groups of the factors are not independent.
SD	Standard deviation
SEM	Standard error of the mean
SIPI	Short Interpulse Interval: Describes a special pair of two pulses in a CI pulse train. The term “short“ is not finally quantified in this context.
THR	Threshold: Lowest current at which a signal is detected by the participant.