DISSERTATION | DOCTORAL THESIS

Titel | Title

Towards Improving Selective Hearing in Cochlear-Implant Listeners

verfasst von | submitted by

Dipl.-Ing. Martin Julius Lindenbeck BSc

angestrebter akademischer Grad | in partial fulfilment of the requirements for the degree of

Doktor der Naturwissenschaften (Dr.rer.nat.)

Wien | Vienna, 2025

Studienkennzahl lt. Studienblatt | Degree programme code as it appears on the student record sheet:	UA 796 605 298
Dissertationsgebiet lt. Studienblatt | Field of study as it appears on the student record sheet:	Psychologie
Betreut von | Supervisor:	Mag. Dr. Bernhard Laback Privatdoz. Univ.-Prof. Dr. Ulrich Ansorge

This page is intentionally left blank

Abstract

Selective hearing, the ability to focus on a desired sound source amidst competing sound sources, is significantly impaired in cochlear-implant (CI) users compared to individuals with normal hearing (NH). Current CI systems, while effective for speech understanding in quiet environments, struggle to replicate the nuanced NH timing cues, particularly temporal pitch and interaural time difference (ITD), required for selective hearing in noisy or complex acoustic settings. These deficits stem both from constraints of the CI electrode-neuron interface and ineffective encoding of timing cues in speech-focused CI stimulation strategies. This doctoral thesis investigates basic approaches towards enhancing selective hearing in CI listeners by addressing limitations in access to salient timing cues.

Four psychophysical studies with human CI listeners were conducted to explore stimulation approaches aimed at improving timing-cue perception. The first study demonstrated that inserting short inter-pulse interval (SIPI) pulses into amplitude-modulated (AM) high-rate pulse trains significantly improves temporal-pitch sensitivity, particularly at low AM depths. The second study revealed that SIPI-based cues generally yield higher perceptual salience than AM-based cues. The third study showed that channel interactions in dual-electrode low-rate stimulation degrade timing sensitivity, depending on between-electrode asynchrony and electrode separation. The similarity of effects for unilateral (temporal pitch) and bilateral (ITD) stimulation suggest a monaural peripheral origin. The fourth study, focusing on unilateral temporal-pitch sensitivity, revealed that the detrimental between-electrode asynchrony effects are similar for low-rate, high-rate AM, and SIPI stimulation. Analysis of modeled auditory nerve responses suggests that channel interaction-induced increases in effective rate and coarser AM coding are the causes underlying the asynchrony effects.

Taken together, this thesis provides basic approaches for improving stimulation approaches, both on the single-electrode level as well as with respect to the across-electrode timing, towards improving timing-cue perception—and thereby a step towards selective hearing—in future CI systems.

Zusammenfassung

Selektives Hören, also das Fokussieren auf eine Zielschallquelle in Anwesenheit von Störschallquellen, ist bei Nutzer:innen von Cochlea-Implantaten (CIs) verglichen mit Normalhörenden (NH) erheblich beeinträchtigt. Aktuelle CI-Systeme erlauben Sprachverständnis in Ruhe. Sie haben jedoch Schwierigkeiten, die nuancierten NH-Reizmerkmale, insbesondere zeitliche Tonhöhe und interaurale Laufzeitdifferenz (ITD), zugänglich zu machen, die für selektives Hören erforderlich sind. Diese Dissertation untersucht grundlegende Ansätze zur Verbesserung des selektiven Hörens bei CI-Hörer:innen, indem sie Einschränkungen beim Zugang zu salienten zeitlichen Reizmerkmalen adressiert.

Vier psychophysikalische Studien mit CI-Hörer:innen wurden durchgeführt. Die erste Studie zeigte, dass das periodische Einfügen von Pulsen mit kurzen Inter-Puls-Intervallen (SIPI) in amplitudenmodulierte (AM) hochratige Pulsketten die zeitliche Tonhöhensensitivität signifikant verbessert, insbesondere bei niedrigen AM-Tiefen. Die zweite Studie zeigt, dass SIPI-basierte Tonhöhe eine höhere perzeptive Salienz aufweist als AM-basierte Tonhöhe. Die dritte Studie zeigte, dass Kanalinteraktionen zwischen Elektroden bei Zweielektrodenstimulation mit niedriger Pulsrate die zeitliche Sensitivität in Abhängigkeit der Elektrodenasynchronität und des Elektrodenabstands beeinträchtigen. Die Ähnlichkeit der Resultate bei unilateraler (zeitliche Tonhöhe) und bilateraler (ITD) Stimulation zeigt Evidenz für einen peripheren Ursprung der Beeinträchtigung. Die vierte Studie zeigte starke Übereinstimmung der nachteiligen Effekte der Asynchronität zwischen den Elektroden auf die unilaterale zeitliche Tonhöhensensitivität für Stimulation mit niedriger Rate, hoher Rate und AM, oder SIPI. Eine Analyse modellierter Hörnervenantworten deutet auf eine durch Kanalinteraktionen induzierte Erhöhung der effektiven Stimulationsrate sowie schlechterer AM-Kodierung als Ursachen hin.

Zusammengenommen bietet diese Dissertation grundlegende Ansätze zur Verbesserung von Stimulationsstrategien, sowohl auf Einzelelektrodenebene als auch elektrodenübergreifend, hinsichtlich der Verbesserung der Wahrnehmung zeitlicher Reizmerkmale und damit als ein Schritt hin zu selektivem Hören mit CIs.

ACKNOWLEDGMENTS

This dissertation has accompanied me through a significant chapter of my adult life and I would like to entend my gratitude to all those who have supported me in various ways along the journey.

First and foremost, I wish to express my sincere appreciation to my supervisor, Bernhard Laback, whose guidance, steadfast support, and personal friendship have been instrumental in shaping both this dissertation and my personal growth.

I am deeply grateful to my co-supervisor, Ulrich Ansorge, as well as my co-authors, Piotr Majdak and Sridhar Srinivasan. Their expertise and insights were particularly valuable during the early stages of my doctoral studies and throughout the research design process.

My appreciation also goes to my (former) colleagues at the Acoustics Research Institute of the Austrian Academy of Sciences, especially Michael Mihocic and Lisa Frohmann, for their assistance in the laboratory and for the many friendships that have developed beyond the workplace.

On a more personal note, I am profoundly grateful to my parents, Elisabeth and Rainer, for their persistent encouragement and unconditional support throughout an academic journey spanning more than twelve years. Your belief in the value of higher education has been fundamental to this achievement.

On an even more personal level, I wish to express my deepest gratitude to my own small family, Marie and Laurin, for their unwavering support and for continually providing both intellectual and emotional encouragement. You have shown me every day what is at least as important as personal achievements——namely, you.

Lastly, I thank all those not mentioned here by name whose contributions have nonetheless been pivotal in helping me reach this milestone.

This page is intentionally left blank

Contents

This page is intentionally left blank

General Introduction

Selective hearing refers to the remarkable ability of the human auditory system to decompose a mixture of concurrent sounds arriving at the two ears into the constituting sources and, subsequently, to focus on one source while ignoring the others. For example, in so-called "cocktail-party" situations, listeners with normal hearing (NH) are able to selectively attend to a target speaker against a background of competing speakers.

The cochlea is a major processing stage for selective hearing along the auditory pathway. In NH, the basilar membrane mechanically decomposes the acoustic signal into its frequency components, a frequency-to-place conversion called tonotopy. Subsequently, inner hair cells convert the frequency components into neural activity by eliciting spikes in spiral ganglion cells of the auditory nerve (AN). Therefore, any place-specific AN activity *per se* (i.e., regardless of the temporal pattern of AN activity) provides so-called place cues for the auditory system. Within an AN fiber, however, the acoustic frequency can furthermore be encoded temporally in the pattern of AN activity via the so-called phase-locking property of AN fibers. Phase-locked AN activity is the basis for temporal cues in the auditory system.

Cochlear implants (CIs) are currently the most widely applied auditory prostheses in cases of severe to profound cochlear hearing loss or deafness. They directly stimulate the remaining spiral ganglion neurons of the AN with an array of intracochlear electrodes emitting stimulation pulses into the cochlear fluid and broadly towards the AN. Despite great success, the performance of CI users, that is, with so-called electric hearing, is still much worse than that of NH listeners, both in controlled laboratory environments using research interfaces and with clinical CI signal processors in everyday life. In quiet, speech understanding is generally good, yet CI users show major deficits in understanding speech in noisy environments and in selective-hearing tasks. Beyond their main purpose of rehabilitation, CIs are important for basic hearing research. In NH, temporal and place cues often co-vary. With CIs, these cues can, with certain restrictions, be varied independently. Hence, CI research can complement NH research and, by contrasting CI with NH performance, aid our general understanding of the human auditory system. In particular, CIs are a powerful means for studying temporal

cues. Beyond that, any selective-hearing capability with current CIs probably largely relies on temporal cues, in particular, temporal pitch and interaural time difference (ITD).

The most important peripheral limitation of electric hearing with current CIs is the electrode-neuron interface (ENI). It describes the *functional* coupling between CI electrodes and AN fibers and defines the fidelity of auditory information transmission. ENI quality depends on factors such as the distance between electrode and AN, the current spread through the intracochlear fluid, and the survival or health of the AN fiber, respectively. Unwanted interactions between stimulation pulses emitted in rapid succession by proximal CI electrodes can degrade both temporal and place cues. Temporal cues are smeared by pulse interactions at the level of the spiral ganglion cell—so-called channel interactions—, occurring even in the absence of electric field interactions. Such interactions may in turn result in unwanted AN activity or suppression, thereby also smearing place cues. Channel interactions are generally thought to increase with increasing stimulation pulse rate.

Selective hearing in cocktail party settings requires both salient speech decoding and source-segregation, in particular timing, cues. However, with current CIs, there is a two-dimensional trade-off between the salient coding of speech and timing cues. The first dimension is the pulse rate: speech is best coded at higher rates that ensure sampling of brief events such as plosives while timing is best coded at lower rates that ensure salient pitch and ITD cues. The second dimension is the number of active electrodes (i.e., the bandwidth): speech requires several active electrodes, for example, to convey vowel formants, while timing of a single source is best coded with only a single active electrode.[1] Crucially, as noted above, both dimensions moderate the amount of channel interactions as well.

The goal of this thesis is to further our basic understanding of timing-cue perception in CI listeners and under controlled laboratory conditions, that is, with research interfaces and stimulating only one or two electrodes per ear. Such improved understanding is pivotal for future improvements in multi-channel CI stimulation that should allow more selective hearing. To that end, in this thesis both single-electrode temporal coding, particularly pulse, AM rate,

[1] In general, selective hearing requires processing of two or more sources. But, at a given point in time, it may be sufficient to code only one source (cf. Gaudrain and Carlyon, 2013; Schoenmaker and van de Par, 2016).

and combination of AM and pulse rate, as well as dual-electrode effects, particularly channel interactions, were investigated in four empirical studies that are outlined below. Throughout this thesis, the best possible prerequisites for selective hearing are assumed to be a combination of maximum single-electrode sensitivity, minimum channel interactions, and as little (pulse-rate) trade-off as possible between timing cues and speech.

Physiology of the Electrically Stimulated Cochlea

In the electrically stimulated cochlea, there is some distance between CI electrode array and AN fibers. Furthermore, the array is located in an electrically conductive fluid. Hence, the number of independent cochlear stimulation sites is limited by the spread of the electric fields of both adjacent and more distant electrodes (e.g., Fu and Nogaki, 2005). Interactions can be due to (1) simultaneous activation of electrodes (electric-field interactions), (2) "delayed" summation of non-simultaneous electrical charges (facilitation effect), or (3) due to neural masking (refractory effect, e.g., de Balthasar et al., 2003). Hereby, electric-field interactions are most detrimental. Hence, CI electrodes do not stimulate continuously but pulsatile and interleaved across electrodes such that—in most cases—only one electrode is active at a time. For clinical safety reasons, pulses are biphasic and charge-balanced.

The CI electrode array allows to grossly reproduce the tonotopic stimulation pattern of the acoustically stimulated cochlea. Place cues are conveyed by the electrode position and temporal cues are conveyed by the electrode stimulation pattern. Still, AN firing patterns differ in multiple ways between acoustic and electric stimulation. When stimulated acoustically, AN fibers have a greater dynamic range, a more variable firing rate, weaker within-fiber phase-locking, and more independent phase-locking across fibers. In electric hearing, AN fibers show strong phase-locking to isochronous electric equal-amplitude pulses presented at low rates or peaks of the amplitude modulation superimposed on isochronous pulses presented at high rates (e.g., Hancock et al., 2017). For a review of electric stimulation of spiral ganglion neurons, see Boulet et al. (2016).

For single-*pulse* stimulation, both positive (anodic) and negative (cathodic) phase can evoke spikes, yet, with different efficiency and different latency (i.e., lower latencies are

associated with positive-phase spiking and central [vs. peripheral] activation of AN fibers). Phase sensitivity has been shown to be opposite in animals (e.g., cats) and humans, with animals showing higher sensitivity for negative phases and humans for positive phases (e.g., Macherey et al., 2008). Therefore, since the differences between animals and humans in the phase-latency relationship have not been fully understood, recent models of electrically stimulated AN fibers differ in whether they account for polarity differences (e.g., Joshi et al., 2017) or not (e.g., Horne et al., 2016). Furthermore, a notable difference is their ability to quantitatively reproduce neural activity in response to different pulse shapes (e.g., monophasic vs. biphasic; Horne et al., 2016; Joshi et al., 2017; Takanen and Seeber, 2022).

For stimulation with more than one pulse, Boulet et al. (2016) note two temporal stimulus-response phenomena which are determined by the properties of spiral ganglion neurons, that is, membrane capacitance and types of voltage-gated ion channels, and are primarily relevant for this project. Each of the phenomena operates in a different time range and therefore at different pulse rates. First, refractoriness (0 to $\sim$ 1 ms) shows as reduced excitability for the second pulse, given a spike in response to the first pulse. After an absolute refractory period, the probability of a spike in response to the second pulse increases exponentially (e.g., Miller et al., 2001). Second, facilitation (temporal summation) refers to the integration of two subsequent pulses with small inter-pulse intervals (IPIs) by the neural membrane, evoking a spike in response to the second pulse. It typically occurs between 0.1 to 0.3 ms (Cartee et al., 2000, 2006). However, some studies suggest that this can extend up to 0.6 ms (Karg et al., 2013). Facilitation effects are stronger for pulse pairs with alternating polarity (Karg et al., 2013) and crucially depend on both the stimulus level and the pulse rate (Heffer et al., 2010). In general, both refractoriness and facilitation are often causing undesired temporal interactions in the electrically stimulated AN, particularly (1) when stimulating with high pulse rates and/or (2) when occurring between stimulation pulses emitted from proximal electrodes (cf., e.g., Middlebrooks, 2004). Recent models of electrically stimulated animal AN fibers have replicated both refractory and facilitatory effects for pulse-train stimulation, however, varying in the scope of successfully modeled data and computational complexity (Felsheim & Dietz, 2024; Joshi et al., 2017; Takanen & Seeber, 2022).

Psychophysics of Electric Stimulation

Speech

Speech understanding with CIs and in quiet is assumed to depend primarily on information encoded in the slowly varying temporal envelope of the signal (Smith et al., 2002; Zeng et al., 2004). Yet, NH listeners use the rapidly varying temporal fine structure of the signal to understand speech in noise (e.g., Lorenzi et al., 2006). Because in the development the primary goal of multi-channel CIs was to restore basic speech understanding in quiet, clinical stimulation strategies—algorithms that transform the acoustic signal into electrode signals—typically focus on transmitting temporal envelope information. Continuous interleaved sampling (CIS, Wilson et al., 1991) is the cornerstone of envelope-based CI stimulation paradigms. CIS splits the acoustic signal into frequency bands (channels) which are allocated to tonotopically matched electrodes. Within each channel, the compressed envelope signal is superimposed on periodic high-rate pulse trains by means of amplitude modulation (AM). Electric-field summation is greatly reduced by stimulating only one electrode at a time. The periodic pulse trains form the electric temporal fine structure. Contrary to the acoustic temporal fine structure, the electric fine structure in CIS-based strategies does not convey information of the acoustic input signal.

For CIS-based strategies, high speech understanding requires sufficiently high pulse rates, that is, at least 500 to 800 pulses per second (pps; e.g., Arora et al., 2009; Friesen et al., 2005; Loizou et al., 2000). Regarding the spectral representation, recognition scores for vowels and consonants improve up to a number of eight electrodes to high but worse-than-NH scores (Friesen et al., 2001). Yet, recent evidence suggests that spectral resolution might saturate beyond 12-channel configurations (Croghan et al., 2017).

To improve speech understanding in noise, electric temporal fine structure in the form of low and slowly varying pulse rate (i.e., low-rate frequency modulation) has been proposed (e.g., Zeng et al., 2005). Several stimulation paradigms have been proposed that aim to incorporate this approach. Examples are fundamental asynchronous stimulus timing (FAST, Smith, 2010), peak-derived timing (PDT, van Hoesel et al., 2008), and fine structure

processing (FSP, Hochmair et al., 2006; FS4/FS4-p, Riss et al., 2014). Improvements in speech understanding in noise have so far been limited (e.g., Dillon et al., 2016; Riss et al., 2011; Zirn et al., 2016) and have furthermore not been associated with electric temporal fine structure (Zirn et al., 2016).

Pitch

Both periodicity and cochlear place can elicit a pitch percept (temporal and place/spectral pitch, respectively). For NH, it was shown that pitch is more salient when conveyed via the temporal fine structure rather than the temporal envelope (Smith et al., 2002; Zeng et al., 2004). Oxenham et al. (2011) suggested that the strong pitch associated with resolved harmonics might reflect place rather than temporal fine-structure cues, although above-chance discrimination of three simultaneous tones is possible based only on temporal cues (Graves & Oxenham, 2019).

Pitch perception in electric hearing is substantially altered compared to NH. Although CI listeners are generally able to perceive both temporal and place pitch, their ability to discriminate temporal pitch in the form of pulse rate differences is limited to approximately 300 pps (e.g., Ihlefeld et al., 2015), with some exception of star listeners (e.g., Kong and Carlyon, 2010; Townshend et al., 1987). CI listeners are furthermore sensitive to temporal-envelope pitch if the pulse rate is sufficiently high (e.g., McKay et al., 1994), with best sensitivity at AM rates of 100 Hz (e.g., Chatterjee and Oberzut, 2011), and limited to AM rates below about 300 Hz (e.g., Kong et al., 2009) similar to pulse rate pitch. Above 300 pps or Hz, respectively, place pitch is more dominant than temporal pitch (Zeng, 2002). Note, however, that the upper limit of temporal pitch may depend on the type of pitch cue (i.e., pulse vs AM rate; Goldsworthy et al., 2022), the CI manufacturer (e.g., Kong et al., 2009) and, crucially, the individual CI listener (e.g., De Groote et al., 2024). Because CI electrodes are static in the cochlea, an increase in acoustic frequency does not necessarily result in a proportional change in place of excitation. Some studies have argued that such a rate-place co-variation is important for salient pitch perception (e.g., Carlyon and Deeks, 2002).

In clinical practice, multi-electrode CIS-like stimulation is most common. Some

studies have investigated multi-electrode temporal-envelope pitch sensitivity. Slight benefits of multi-electrode stimulation compared to single-electrode configurations were found (e.g., Geurts and Wouters, 2001), but probably mostly due to increased overall loudness (Galvin et al., 2015). Importantly, multi-electrode place pitch is rather weak (e.g., Laneau and Wouters, 2004). Previous attempts to improve temporal-envelope pitch coding in CI stimulation paradigms were mostly aimed at increasing the AM depth (e.g., F0mod; Laneau et al., 2006; Milczynski et al., 2009) or improving the AM shape (e.g., sawtooth modulation; Green et al., 2004, 2005), so far with varying degrees of success.

Interaural Differences in Time and Level

Binaural cues, that is, ITD and interaural level difference (ILD), are crucial for horizontal-plane sound localization (e.g., Macpherson and Middlebrooks, 2002) and speech recognition in stationary noise (Bronkhorst & Plomp, 1988). In addition to that, ITDs are important for spatial release from speech-in-speech masking (e.g., Kidd et al., 2010). For frequencies up to 1.4 kHz (Brughera et al., 2013; Zwislocki & Feldman, 1956), fine-structure ITD cues dominate horizontal-plane sound localization whereas above, ILD cues dominate and envelope-ITD cues contribute to some extent (e.g.,Klingel and Laback, 2022; Macpherson and Middlebrooks, 2002). Reliable real-world ITD cues are encoded in the rising portions (i.e., the onset) of the AM (Dietz et al., 2013).

In electric hearing, phase-locking is accurate for frequencies up to 1 kHz (Dynes & Delgutte, 1992). Yet, using direct stimulation of electrodes via a research interface, behavioral sensitivity to pulse-rate ITD degrades for rates exceeding 100 pps and become immeasurable between 400 and 800 pps (e.g., Laback et al., 2007; van Hoesel et al., 2009), broadly similar to pulse rate pitch (e.g., Ihlefeld et al., 2015). At best rates around 100 pps, ITD discrimination thresholds are an order of magnitude worse than the best NH thresholds (see, e.g., Laback et al., 2015).

CI listeners are overall sensitive to ITDs in the temporal envelope, albeit much less than NH listeners (e.g., Majdak et al., 2006). Furthermore, the envelope shape (i.e., AM depth, onset steepness, and off-time) crucially affects ITD senstivitiy (e.g, Laback et al., 2011;

Monaghan and Seeber, 2016). There appears to be a similarity in the rate or frequency dependence, respectively, between fine structure and envelope ITD in electric hearing (Noel & Eddington, 2013; van Hoesel et al., 2009) on the one side and envelope ITD in NH on the other side (L. R. Bernstein & Trahiotis, 2002). Contrary to NH, where the tonotopic place of stimulation is crucial for ITD sensitivity, CI thresholds do not vary substantially with the place of stimulation (Laback et al., 2015). Recent studies on laboratory-controlled multi-electrode fine-structure ITD found broadly constant sensitivity for multiple interaural electrode pairs compared to single pairs (Egger et al., 2016; Ihlefeld et al., 2014; Kan et al., 2015). Multi-electrode mixed-rate approaches—employing low-rate stimulation at one electrode in basal or middle cochlear regions among four electrodes with high-rate stimulation—yielded low-rate-like ITD sensitivity (Thakkar et al., 2018, 2023).

As a result of ear-specific differences in CI insertion depth, interaural stimulation site mismatches occur. Typical mismatches result in degraded ITD and ILD sensitivity, distorted spatial maps, multiple or non-centered auditory images, and impaired lateralization (Goupell, Kan, & Litovsky, 2013; Goupell, Stoelb, et al., 2013; Kan et al., 2013). Psychophysically, interaural mismatches have often been estimated by matching place pitch between ears (e.g., Kan et al., 2013; Laback et al., 2007; van Hoesel, 2004). Yet, more recent evidence suggests that ITD-based rather than place-pitch-based matching best aligns with both brainstem responses (Hu & Dietz, 2015) and computed-tomography scans (J. G. Bernstein et al., 2021).

The CI stimulation paradigm crucially determines binaural sensitivity. With CIS-like stimulation, only envelope ITD cues are available. For speech stimuli, ITD sensitivity is rather weak (Grantham et al., 2008; Laback et al., 2004). CI listeners thus rely on ILD cues in sound lateralization (Laback et al., 2004) as well as in localization of sound sources (e.g. Kerber and Seeber, 2013). Stimulation paradigms developed to provide cues via the electric temporal fine structure (cf. Speech section) have so far enhanced ITD sensitivity only for narrowband but not broadband acoustic stimuli (Fischer et al., 2021; Zirn et al., 2016), pointing towards detrimental effects of channel interaction and/or across-electrode timing.

One proposed approach to reduce the trade-off between ITD and speech perception, that is, to improve ITD sensitivity at high pulse rates, is to apply binaurally coherent jitter to

the IPIs of pulse trains (Laback & Majdak, 2008). Single-electrode ITD sensitivity was shown to improve for rates high enough to encode speech. Neurophysiological data showed that the improvements arise from randomly occurring very short IPIs (Hancock et al., 2012). Such perceptually relevant short IPIs will hereafter be referred to as SIPIs.

The idea of selectively enhancing ITD sensitivity at high rates using a deterministic SIPI approach instead of a random jitter-based approach, by inserting SIPI pulses at a certain (SIPI) rate, was confirmed in both psychophysical (Srinivasan et al., 2018, 2020) and neurophysiological experiments (Buechel et al., 2018) for single-channel stimulation. For unmodulated 1000-pps pulse trains, SIPI pulses yielded jitter-like behavioral ITD sensitivity for SIPI rates up to 200 pps (Srinivasan et al., 2018). For these SIPI rates, ITD sensitivity was furthermore at low-rate level (i.e., pulse trains with a rate equal to the SIPI rate in the 1000-pps pulse trains). For 1000-pps with an AM of either 125 or 250 Hz and SIPIs placed at AM peaks at a rate of 63 Hz, performance improved significantly for both AM rates and across all tested AM depths compared to AM-only pulse trains (Srinivasan et al., 2020).

Research Questions and Thesis Outline

In four studies, this thesis is concerned with investigating the perception of temporal pitch and ITD—hereafter collectively referred to as timing cues—using direct electrical stimulation of either one or two CI ears and, per ear, one or two electrodes. Two major dimensions of selective hearing with CIs were at the center of the investigations, in particular

1. the stimulation approach, that is, the general per-electrode pulse pattern, crucially determining timing-cue salience and speech-encoding capability, and

2. the across-electrode stimulation timing, crucially determining the amount of channel interactions and potentially interaction with the per-electrode stimulation approach.

Primarily psychophysical methods and, in the last study, complementary phenomenological modeling of the electrically stimulated AN were employed. Broadly reflecting voiced-speech coding with CIS, the rate at which timing cues were conveyed is referred to as the fundamental frequency (F0) throughout the thesis.

In study one (Lindenbeck et al., 2020), based on the promising results of Srinivasan

et al. (2018, 2020), we employed the SIPI approach aiming to improve temporal-pitch sensitivity for single-electrode stimulation. To get a comprehensive overview across stimulation approaches, we compared SIPI stimulation to unmodulated low-rate pulse trains (LR stimuli) and high-rate pulse trains with an AM but without SIPI pulses (HR stimuli). Hereby, the F0 corresponded either to the pulse rate (LR stimulation), the SIPI rate (SIPI stimulation) and/or the AM rate (SIPI and HR stimulation). The main parameters were the F0, the AM depth, the carrier rate, and the ratio of SIPI rate and F0.

In study two (Lindenbeck, Majdak, et al., 2023), following up on open questions from study one and to investigate whether the SIPI approach can be used to jointly enhance temporal-pitch and ITD sensitivity, we investigated the relative perceptual salience of SIPI-based and AM-based temporal pitch with single-electrode stimulation and for SIPI rates low enough to potentially enhance ITD sensitivity (cf. Srinivasan et al., 2020). The main parameters were the SIPI and AM rate, the AM depth, and consistency of SIPI and AM cue.

In study three (Lindenbeck et al., 2024), we investigated across-electrode timing with dual-electrode LR stimulation for both unilateral stimulation (testing temporal-pitch sensitivity) and bilateral stimulation (testing ITD sensitivity) to investigate effects and origin of channel interactions for timing-cue perception. The main parameters were the monaural temporal asynchrony, tonotopic separation of electrodes, and the percept.

In study four (Lindenbeck et al., 2025), we focused again on unilateral dual-electrode stimulation and temporal pitch to investigate the effect of stimulation approach on channel interactions. The main parameters were the stimulus type, the monaural temporal electrode asynchrony, and the tonotopic separation. To comprehensively discuss the results in light of the potential cues provided by the different stimulus types, we used a phenomenological model of the AN to generate spiking patterns and, subsequenctly, IPI distributions as indicators of instantaneous rate coding.

Across studies, we consider the beneficial combination of stimulation approach and across-electrode timing to pave the way towards more selective hearing with CIs in future stimulation strategies.

Empirical Work

Temporal-Pitch Sensitivity with Pulse Rate, Amplitude Modulation, and Short Inter-pulse Interval Cues

The following manuscript was published in the Psychological and Physiological Acoustics section of the Journal of the Acoustical Society of America[2]:

Lindenbeck, M. J., Laback, B., Majdak, P., & Srinivasan, S. (2020). Temporal-pitch sensitivity in electric hearing with amplitude modulation and inserted pulses with short inter-pulse intervals. *The Journal of the Acoustical Society of America*, *147*(2), 777–793. https://doi.org/10.1121/10.0000610

Author contributions

All authors conceptualized and designed the study. MJL collected the data and performed the statistical analysis. MJL wrote the first draft of the manuscript. All authors contributed to manuscript revision, read and approved the submitted version.

The conceptualization of the study and the collection of data were integral components of my master's thesis. Subsequently, during my doctoral studies, I conducted the published statistical analysis and navigated the entire peer review and publication process.

JASA ARTICLE

Temporal-pitch sensitivity in electric hearing with amplitude modulation and inserted pulses with short inter-pulse intervals[a)]

Martin J. Lindenbeck,[b)] Bernhard Laback,[c)] Piotr Majdak,[d)] and Sridhar Srinivasan[e)]
Acoustics Research Institute, Austrian Academy of Sciences, Wohllebengasse 12-14, 1040 Vienna, Austria

ABSTRACT:
Listeners with cochlear implants (CIs) typically show poor sensitivity to the temporal-envelope pitch of high-rate pulse trains. Sensitivity to interaural time differences improves when adding pulses with short inter-pulse intervals (SIPIs) to high-rate pulse trains. In the current study, monaural temporal-pitch sensitivity with SIPI pulses was investigated for six CI listeners. Amplitude-modulated single-electrode stimuli, representing the coding of the fundamental frequency ($F0$) in the envelope of a high-rate carrier, were used. Two SIPI-insertion approaches, five modulation depths, two typical speech-$F0$s, and two carrier rates were tested. SIPI pulses were inserted either in every amplitude-modulation period (full-rate SIPI) to support the $F0$ cue or in every other amplitude-modulation period (half-rate SIPI) to circumvent a potential rate limitation at higher $F0$s. The results demonstrate that full-rate SIPI pulses improve temporal-pitch sensitivity across $F0$s and particularly at low modulation depths where envelope-pitch cues are weak. The half-rate SIPI pulses did not circumvent the limitation and further increased variability across listeners. Further, no effect of the carrier rate was found. Thus, the SIPI approach appears to be a promising approach to enhance CI listeners' access to temporal-envelope pitch cues at pulse rates used clinically.
 https://doi.org/10.1121/10.0000610

(Received 21 November 2018; revised 26 December 2019; accepted 2 January 2020; published online 5 February 2020)
[Editor: Joshua G. Bernstein] Pages: 777–793

I. INTRODUCTION

Pitch is one of the primary auditory sensations. In music, it is essential for providing a sense of melody and harmony (e.g., Houtsma, 1995). In speech, it is important for conveying prosody (e.g., Cutler *et al.*, 1997) and for lexical-word meaning (Xu and Pfingst, 2003; Chen *et al.*, 2018), speech perception in noise (e.g., Miller *et al.*, 2010), speaker identification, and segregation of simultaneous speakers (e.g., Darwin, 2005). In normal hearing, pitch may be conveyed both by means of the tonotopic place of stimulation along the cochlea, the so-called place code, and by means of the periodicity at a given place of stimulation, the so-called temporal code (e.g., Oxenham, 2012). In electric stimulation of the auditory nerve with cochlear implants (CIs), the temporal code may be more important, because CIs provide poor spectral resolution. The current study was concerned with temporal-pitch perception with CIs.

In acoustic hearing, temporal cues are encoded either via the rapid oscillations related to the carrier frequency (i.e., the acoustic temporal fine-structure) or via the slowly changing amplitude (i.e., the temporal envelope). Fine-structure pitch cues are most salient at lower carrier frequencies whereas envelope pitch cues are most salient at higher carrier frequencies. For example, with harmonic complex tones such as vowels in voiced-speech segments, the fundamental frequency ($F0$) is encoded temporally via the envelope repetition rate of auditory-filter outputs from groups of higher (unresolved) harmonics (residue pitch, Schouten, 1940). The $F0$ may also be derived by combining the temporal fine-structure of low-numbered (resolved) harmonics (Oxenham, 2012). While this pitch is perceived more saliently than the residue pitch, there still remains some debate as to whether it is actually derived from temporal fine-structure and/or place cues (e.g., Oxenham *et al.*, 2011).

CIs are the most successful auditory prostheses for the clinical treatment of severe hearing loss or deafness. A key feature that determines the fidelity of pitch perception in electric hearing is the CI's stimulation paradigm, which defines how the acoustic input signal is converted into electric stimulation sequences. Other major aspects are the design of the implanted electrode arrays as well as biological and listener-specific factors such as, for example, neural survival and cognitive abilities (for a review, see Wilson and Dorman, 2008). Today, most clinically applied CI stimulation strategies focus on transmitting the temporal envelope of the acoustic signal by amplitude-modulating an electric pulse train generated at a constant and high rate. In

[a)]Portions of this work were presented at the 21st Annual Meeting of the German Society for Audiology, Halle (Saale), March 2018, and the 2019 Conference on Implantable Auditory Prostheses, Lake Tahoe, CA, July 2019.
[b)]ORCID: 0000-0001-7779-4767.
[c)]Electronic mail: bernhard.laback@oeaw.ac.at, ORCID: 0000-0003-0929-6787.
[d)]ORCID: 0000-0003-1511-6164.
[e)]ORCID: 0000-0002-6818-7356.

https://doi.org/10.1121/10.0000610 JASA

doing so, the acoustic temporal fine structure is discarded. Thus, the electric temporal fine-structure does not contain information about the acoustic signal.

With current CIs, place-pitch perception is degraded due to the coarse representation of the place code with the few available electrodes and the even lower number of independent spectral channels (e.g., Mehta and Oxenham, 2017). With stimulation strategies not conveying temporal fine-structure information (e.g., continuous interleaved sampling, CIS; Wilson *et al.*, 1991), pitch perception is further restricted to the temporal envelope cues. Pitch coding in the temporal envelope relies critically on the modulation depth of the electrode signals (McKay *et al.*, 1995; Vandali *et al.*, 2013). Because environmental noise and reverberation degrade the effective modulation depth even in normal hearing (e.g., Sayles and Winter, 2008), pitch perception in electric hearing can be expected to be impaired particularly in everyday situations.

Although CI listeners mainly have access to temporal envelope-pitch cues, most data on pitch perception in CI listeners have been collected with unmodulated low-rate pulse trains presented directly to a single CI electrode (bypassing the clinical sound processor). Just-noticeable differences for rate discrimination of roughly a semitone indicate a general ability to perceive pitch in electric hearing (for a review, see Moore and Carlyon, 2005). However, temporal-pitch perception in electric hearing is restricted by a purely temporal limitation, characterized by poor sensitivity to stimuli with rates beyond about 200–300 pulses per second (pps) of the carrier pulses for unmodulated pulse trains (e.g., Townshend *et al.*, 1987; Kong *et al.*, 2009; Vandali *et al.*, 2013), of analog electric sinusoids (e.g., Shannon, 1983; Stupak *et al.*, 2018), and of the modulator of high-rate pulse trains (e.g., Kong *et al.*, 2009). While the neural basis of this limitation is still not fully understood to date, it is noteworthy that there are some exceptional CI listeners showing no indication of a rate limitation for pulse rates up to 900 pps (e.g., Kong and Carlyon, 2010).

Sensitivity to interaural time differences (ITDs) is also limited by pulse rate (e.g., Laback *et al.*, 2007; van Hoesel *et al.*, 2009). Limitations in both ITD and pitch sensitivity involve temporal processing and these limitations appear to at least partially share a common neural mechanism (Ihlefeld *et al.*, 2015). These limitations introduce a trade-off between transmission of speech cues with high fidelity requiring pulse rates of at least 500 (Arora *et al.*, 2009) to 800 pps (Loizou *et al.*, 2000), and transmission of pulse-rate pitch and ITD cues requiring rates below 200–300 pps. One approach to reduce this limitation and to enhance ITD sensitivity at higher carrier rates is to apply a binaurally coherent jitter to the inter-pulse intervals (Laback and Majdak, 2008). Their findings of substantially improved sensitivity were explained from the neurophysiological perspective (Hancock *et al.*, 2012) showing that firing and ITD sensitivity of inferior-colliculus neurons in anesthetized and bilaterally implanted cats was enhanced only at sparse instances of time when the jitter randomly created short inter-pulse intervals (SIPIs). In a follow-up study, Buechel *et al.* (2018) found significantly increased firing and ITD sensitivity with inserted SIPI pulses in about 60% of the inferior-colliculus neurons tested in unanesthetized and bilaterally implanted rabbits. Consistently, in a parallel psychophysical study, Srinivasan *et al.* (2018) showed that the insertion of SIPI pulses into an unmodulated 1000-pps pulse train substantially improved human ITD sensitivity. Most recent psychophysical data showed enhanced ITD sensitivity for SIPI pulses inserted into amplitude-modulated pulse trains (Srinivasan *et al.*, 2020). Both Srinivasan *et al.* (2018) and Buechel *et al.* (2018) showed that an increase in the amplitude of a single pulse provides a similar perceptual benefit as an insertion of a SIPI pulse, suggesting that both methods produce cues of similar magnitude in the neural response.

Motivated by these results for binaural processing and given the similarities in temporal-coding limitations between ITD and rate discrimination, the present study investigated pitch-discrimination sensitivity with inserted SIPI pulses in two experiments. An idealized single-electrode representation of voiced speech, as generated by a typical envelope-based stimulation paradigm, was used. In particular, high-rate amplitude-modulated single-electrode pulse trains (similar to those used in Srinivasan *et al.*, 2020) were used, providing *F*0-pitch cues in the amplitude modulation.

Experiment 1 investigated the effects of modulation depth, stimulus type, and *F*0 on pitch-discrimination sensitivity. A carrier rate of 2000 pps was used, providing sufficient (six-times) oversampling of the maximum envelope rate. *F*0s corresponding to both male and female speech (nominally 125 and 250 Hz, respectively) were tested. Given that the effective modulation depth in everyday listening situations can be largely reduced due to interference from the acoustic environment, the stimulus modulation depth was varied across a wide range. SIPI pulses were inserted at two rates corresponding to either the *F*0 (full-rate-SIPI condition) or half the *F*0 (half-rate-SIPI condition), using the amplitude-modulated pulse train as a reference (no-SIPI condition). The half-rate-SIPI condition was motivated by the idea that, particularly at the high *F*0, temporal-pitch sensitivity in the full-rate-SIPI condition might be affected by the rate limitation for temporal pitch and halving the SIPI rate might circumvent that limitation. Experiment 2 tested whether the effects of modulation depth and stimulus type on pitch-discrimination sensitivity generalized to another carrier rate of 1000 pps. Here, only the low nominal *F*0 of 125 Hz was tested.

SIPI pulses inserted into amplitude-modulated high-rate pulse trains are assumed to introduce transient (and therefore temporally precise) timing information even in neurons that are otherwise in a rate-adapted state (Buechel *et al.*, 2018). The most transient stimulus with CIs is the unmodulated pulse train with low rates, providing steeper slopes and longer off-times than amplitude-modulated high-rate SIPI pulse trains. Hence, unmodulated low-rate single-electrode pulse trains with rates corresponding to the *F*0s and SIPI rates were also tested for comparison with the high-rate conditions and to link the findings to a wider range of literature.

JASA
https://doi.org/10.1121/10.0000610

Experiment 1 addressed different aspects of the basic hypothesis that SIPI pulses inserted into high-rate amplitude-modulated pulse trains would trigger neural spiking events similar to those triggered by unmodulated pulse trains with pulse rates corresponding to the SIPI rate (cf. Srinivasan *et al.*, 2018, 2020; Buechel *et al.*, 2018), thus providing similar temporal-pitch sensitivity. Particularly, the following effects were hypothesized:

(1) For both nominal *F*0s, discrimination performance would be better for the full-rate-SIPI condition than for the no-SIPI condition, especially at low modulation depths where the *F*0 is weakly encoded in the temporal envelope. In the best case, the full-rate-SIPI condition would restore performance to the level observed with unmodulated low-rate stimuli.
(2) At the high *F*0, both no-SIPI and full-rate-SIPI performance would be degraded due to a rate limitation, similar to the effect usually found for low-rate stimuli (e.g., Ihlefeld *et al.*, 2015). Half-rate SIPI pulses would circumvent that rate limitation (if occurring), manifesting in better performance for the half-rate-SIPI condition than for both the no-SIPI and the full-rate-SIPI conditions.

II. EXPERIMENT I: 2000-PPS CARRIER RATE

A. Methods

1. Listeners

Six adult CI listeners (four females) participated in the study, details for whom are listed in Table I. Five of the listeners were post-lingually deafened and one was pre-lingually deafened (CI21). All had 12-channel implants manufactured by MED-EL Corp. (Innsbruck, Austria). Bilateral listeners (all except CI18) were asked to choose their preferred ear with respect to pitch perception. Except for CI21, all listeners' decisions corresponded to the ear that was implanted first. All listeners were paid an hourly wage for their participation.

2. Stimuli and apparatus

Figure 1 illustrates the stimuli for experiment 1. All stimuli were nominally 600 ms long (the figure shows a few cycles only) and were presented to a single electrode. The main stimuli were amplitude-modulated high-rate pulse trains [Figs. 1(A)–1(C)]. The modulation shape was a full-wave rectified sinusoid and was chosen to mimic CIS-like signal processing of voiced speech (e.g., Dorman and Wilson, 2004; Wilson, 2006), as in the corresponding ITD study (Srinivasan *et al.*, 2020). The stimulus modulation rate *after* full-wave rectification defines the *F*0 encoded in the envelope. One experimental parameter was the nominal *F*0, which was defined as the geometric average of the *F*0s to be discriminated (see Sec. II A 3). The tested nominal *F*0s were chosen based on physiological differences across genders, resulting in the average female *F*0 being almost one octave above the average male *F*0 (Titze, 1989). For example, analyses of American-English speakers revealed average *F*0s of 130 and 233 Hz (Peterson and Barney, 1952) or 119 and 210 Hz (Pépiot, 2014) for male and female talkers, respectively. Hence, in this experiment, the low nominal *F*0 was 125 Hz and the high nominal *F*0 was 250 Hz. To ensure sufficient oversampling of all *F*0s occurring in the experiment (the highest *F*0 was 333 Hz and occurred in the training described in Sec. II A 5), the carrier rate was set to 2000 pps (cf. McKay *et al.*, 1994; Wilson *et al.*, 1997).

Linear 150-ms onset and offset ramps were applied to minimize onset and offset cues. The transitions between the unmodulated ramps and the amplitude-modulated segments were ensured to be smooth. The lengths of the amplitude-modulated segments were adjusted to always contain an integer number of modulation periods which depended on the *F*0. This number was roved by ± 1 per interval to avoid *F*0-specific length cues. Consequently, the stimulus duration, although targeted at 600 ms, depended on the *F*0 and varied by up to ± 10 ms.

Figure 1(A) shows the high-rate stimulus with amplitude modulation only (i.e., the no-SIPI condition). Figure 1(B) shows the high-rate stimulus with SIPI pulses inserted in every amplitude-modulation period, thus with a SIPI rate equal to the *F*0 (full-rate-SIPI condition). Figure 1(C) shows the high-rate stimulus with SIPI pulses inserted in every other amplitude-modulation period, thus with a SIPI rate equal to half of the *F*0 (half-rate-SIPI condition). Previous studies on ITD perception in electric hearing revealed that explicit coding of ITDs might be most effective at peaks of the amplitude modulation (Hu *et al.*, 2017; Srinivasan *et al.*, 2020). Similar

TABLE I. Details for the CI listeners participating in the study. CI18 was unilaterally implanted (right ear only). If the onset of deafness could not be attributed to a specific date (e.g., in case of progressive hearing loss), "age at onset of deafness" may also refer to the onset of profound hearing loss.

				Age at onset of deafness (years)		Ear used for the experiment		
Listener	Gender	Age at testing (years)	Etiology	L	R	Side	CI experience (yr)	Implant type
CI17	female	71	Idiopathic	41	41	R	13	Synchrony
CI18	male	60	Progressive	-	47	R	13	Pulsar
CI21	female	24	Congenital	0	0	R	19	Pulsar
CI24	female	55	Progressive	40	41	L	13	C40+
CI74	male	48	Progressive	41	41	R	6	Sonata
CI77	female	64	Sudden hearing loss	33	58	R	6	Sonata
Mean	-	53.7	-	31.0	36.2	-	11.7	-

https://doi.org/10.1121/10.0000610

JASA

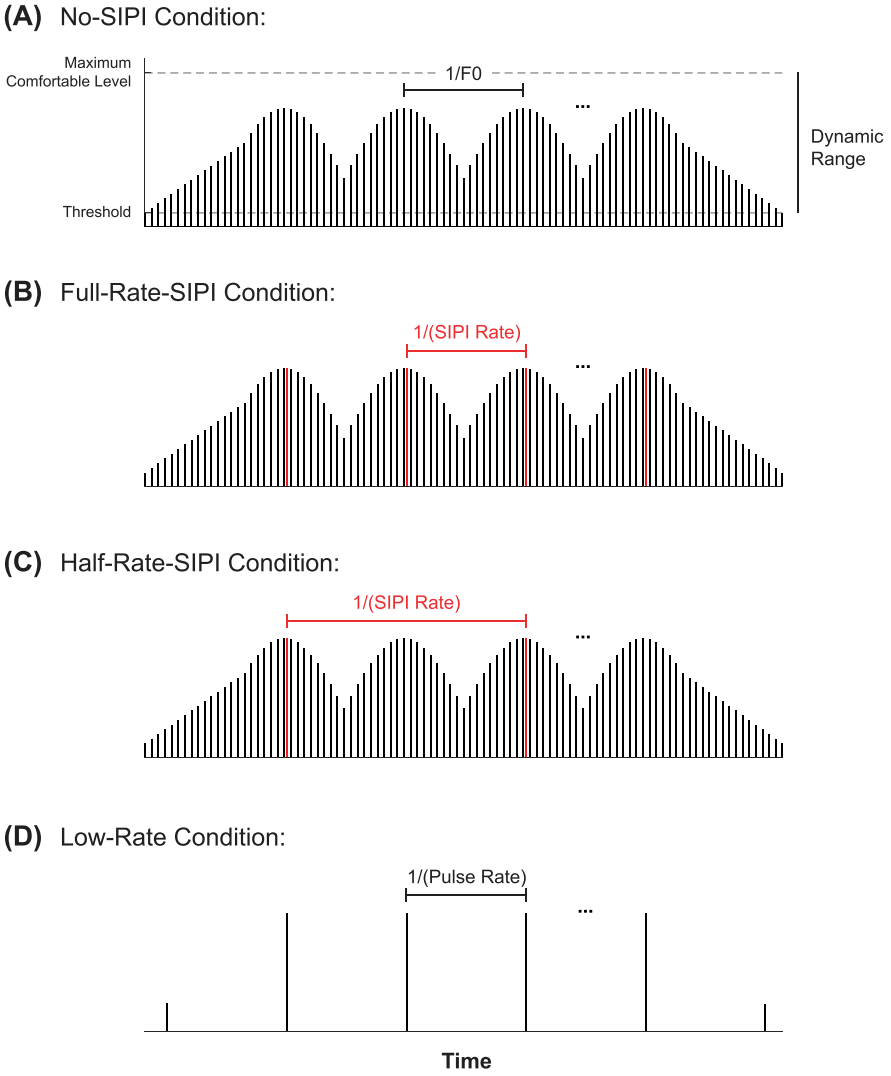


FIG. 1. (Color online) Stimulus conditions: panels (A) to (C) show high-rate pulse trains, amplitude modulated with a full-wave rectified sinusoid. The temporal envelope encodes the fundamental frequency ($F0$). For readability purposes, the negative phases of the biphasic pulses are omitted and only a few amplitude modulation cycles of the actual signals are shown. Panel (A), no-SIPI condition: the pulse train is purely amplitude modulated. Threshold and maximum comfortable level together define the electric dynamic range. Panel (B), full-rate-SIPI condition: a SIPI pulse is inserted at every envelope peak, thus SIPI rate and $F0$ are identical. Panel (C), half-rate-SIPI condition: a SIPI pulse is inserted at every other envelope peak, thus the SIPI rate is half the $F0$. Panel (D), low-rate condition: unmodulated pulse trains, the pulse rate corresponds to $F0$, SIPI rate, or both.

results were obtained from the inferior colliculus of anesthetized cats with CIs showing precise phase locking mostly to modulation peaks (Hancock *et al.*, 2017). Thus, for all high-rate conditions, a carrier pulse was always placed exactly at the peak and a SIPI pulse was always inserted after that carrier pulse. The SIPI fraction (cf. Srinivasan *et al.*, 2018), defined as the interval between the regular and the SIPI pulse as a fraction of the carrier period, was set to 12%. This resulted in a time gap of 60 μs between the regular and the SIPI pulse. Note that SIPI pulses were not inserted in the ramped segments of the stimuli.

Figure 1(D) shows the stimulus for the low-rate condition. Pulse rates were 62.5, 125, and 250 pps, covering all $F0$s and SIPI rates tested. The pulse properties for the low-rate stimuli were otherwise identical to those for the high-rate stimuli.

All pulses were symmetric, biphasic, and cathodic-phase leading. The duration of a phase was 26.7 μs and there was no inter-phase gap. For CI21, the phase duration had to be increased to 33.3 μs to reach the maximum comfortable level [cf. Fig. 1(A)] in the fitting procedure. Consequently, for CI21, the SIPI fraction had to be increased to 14%. The implants stimulated in monopolar mode.

Electrode 8 was used for all listeners. For the implants used by the listeners in this study, this electrode is located slightly more basally than the location corresponding to the carrier rate in acoustic hearing, assuming a typical cochlear length and full electrode insertion (Baumann and Nobbe, 2004), and this choice of electrode thus provided approximately matched carrier-rate and place-pitch cues.

The stimulus amplitudes were defined in linear units relative to the dynamic range [see Fig. 1(A)]. To estimate the dynamic range for each listener, threshold and maximum comfortable level for electrode 8 were manually determined using a fitting procedure with unmodulated 2000-pps pulse trains.

The stimuli were created on a personal computer and sent directly to the implants via the Research Interface Box II, developed at the Institute of Ion Physics and Applied Physics, Leopold-Franzens-University of Innsbruck,

JASA
https://doi.org/10.1121/10.0000610

Austria. The research interface allowed the precise control of the pulse timing and the bypassing of the clinical sound processor.

3. Tasks

Listeners were tested using two-interval, two-alternative forced-choice (2I-2AFC) tasks. Within each trial, stimuli were separated by a gap of 490 ms. In the loudness-balancing pretest (see Sec. II A 4), listeners had to indicate which of the two intervals contained the louder stimulus. In the pitch-discrimination tests (see Secs. II A 5 and II A 6), the listeners had to indicate whether the stimulus in the second interval was higher or lower in pitch than the stimulus in the first interval. All listeners were further instructed to focus on the dominant pitch in case they heard more than one pitch. A double-pitch percept was predicted to occur in conditions having differing *F*0s and SIPI rates (i.e., in the half-rate-SIPI condition).

Responses were provided with a hand-held controller. After each response, listeners were given feedback to reduce response bias (Klein, 2001) and a new trial was played after a break of 310 ms. If a listener felt uncomfortable with the stimulus timing, this break was adjusted. Before each new task, listeners were orally instructed. Written instructions were available at all times.

4. Loudness balancing

The insertion of SIPI pulses into unmodulated high-rate pulse trains can increase the loudness and this effect grows with increasing SIPI rate (Srinivasan *et al.*, 2018). Also, the modulation depth can affect loudness (McKay and Henshall, 2009). At the same time, pitch sensitivity may depend on loudness (Carlyon *et al.*, 2010). To avoid loudness as a potential confound in this experiment, the stimuli for all experimental conditions were loudness-balanced before testing them on pitch. To this end, the amplitudes needed for balanced loudness of all experimental stimuli were measured using an adaptive 3-up 1-down staircase procedure (cf. Jesteadt, 1980). Each trial contained the reference followed by the target, the latter being varied in level. Per condition, pairs of upward and downward staircases with inverse decision rules were run in a randomly interleaved fashion. The upward staircases converged at 79% of louder-perceived targets, the downward staircases converged at 21% of louder-perceived targets. At the beginning of each sequence, target levels were set to 20% of the dynamic range below and above the reference for upward and downward staircases, respectively. The initial step size was 10% of the dynamic range. It was decreased by a factor of 0.7 following each turnaround until it reached its minimum of 2% of the dynamic range. Staircases terminated after 12 turnarounds. The last eight turnarounds were arithmetically averaged to calculate the 79% and 21% levels. Each pair of upward and downward staircase was run twice. If staircases did not converge, they were repeated so that at least four staircases per condition were valid. Per condition, all 79% and 21% levels were arithmetically averaged to calculate the loudness-balanced level (Egger *et al.*, 2017; Srinivasan *et al.*, 2018). Listeners were allowed to take breaks between the runs.

All stimuli in the experiment were loudness-balanced to a single reference, the 125-Hz-*F*0, 0.5-modulation-depth stimulus without SIPI pulses (no-SIPI condition) which was adjusted to have a comfortable level on a categorical loudness scale. The comparison stimuli included targets with all combinations of modulation depth (0.1, 0.3, 0.5, 0.7, and 0.9), *F*0 (125 and 250 Hz), and stimulus type (no-SIPI, full-rate-SIPI, and half-rate-SIPI condition). Note that the no-SIPI condition was loudness balanced only for an *F*0 of 125 Hz because it was assumed that, without SIPI pulses, changes in *F*0 should not evoke changes in loudness (Chatterjee and Oberzut, 2011; Vandali *et al.*, 2013). Additionally, low-rate conditions with pulse rates of 62.5, 125, and 250 pps, corresponding to the nominal *F*0s and SIPI rates were loudness-balanced to the reference.

All participants successfully performed the loudness-balancing procedure. For the full-rate-SIPI and half-rate-SIPI conditions, the amplitudes needed for balanced loudness had to be decreased compared to the no-SIPI condition, particularly at the high *F*0. This is consistent with the increasing SIPI rate associated with increasing *F*0 and can be understood in terms of the pulse-rate effect on loudness for unmodulated pulse trains (McKay *et al.*, 2001).

5. Listener-specific frequency difference

The approach of quantifying sensitivity by means of an individualized frequency difference (FD) as a form of a pitch-discrimination threshold was followed. This allowed for testing a large number of conditions with a single FD, despite the restricted availability of CI listeners (e.g., Ihlefeld *et al.*, 2015). Only this listener-specific FD was used in the pitch-discrimination experiment and its selection aimed at minimizing floor and ceiling effects (see below). The FD was defined in percent relative to the lower of the two F0s.

Performance was quantified by means of the sensitivity index *d'* (Klein, 2001) which controls for response bias towards a certain order of intervals. In the task used in this study, a threshold *d'* score of 1 was equivalent to a bias-free score of 76% correct responses. The maximum performance was set to a *d'* score of 3.3, which was equivalent to a bias-free score of 99% correct responses and prevented the *d'* score from reaching infinity. The nominal *F*0 was defined as the geometric average of lower and higher *F*0s within a trial (cf. Kreft *et al.*, 2010). Note that, due to the definition of the SIPI fraction, the *F*0s were allowed to be only integer submultiples of the carrier rate. Thus, the actual average *F*0s used in the experiments slightly deviated from the targeted nominal *F*0, restricting the range of realizable FDs.

Changing the *F*0 between the intervals of a trial might have resulted in loudness differences potentially evoking confounding cues in the pitch-discrimination task. Although

https://doi.org/10.1121/10.0000610 JASA

the *F*0 encoded in the envelope of a high-rate carrier pulse train does not seem to have an effect on loudness (Chatterjee and Oberzut, 2011), potential subtle residual loudness differences were minimized by applying level roving in the range of 3% of the dynamic range on each interval. However, this might have had an effect on overall pitch sensitivity (Carlyon *et al.*, 2010).

Prior to the FD estimation, listeners were trained on the pitch-discrimination task to reduce task-learning effects. The training was done using FDs of 67% and 100% for the low nominal *F*0 and of 83% for the high nominal *F*0. The modulation depth was 0.3 and both no-SIPI and full-rate-SIPI conditions were included. The method of constant stimuli was applied and all conditions were tested 100 times. Per condition, in half of the trials, the first interval contained the lower *F*0. In the other half of the trails, the first interval contained the higher *F*0. All conditions were tested together in one block in a randomized order of trials. Participants were able to take breaks whenever they wanted. After the training, all listeners (except CI24 in the no-SIPI condition) were able to discriminate these pitch differences.

The subsequent FD estimation was done in separate blocks for the two nominal *F*0s. For each listener, an FD for which the sensitivity across various experimental conditions was distributed across the entire performance range (i.e., for which the performance in the different conditions varied well between *d'* scores of 0 and 3.3 without showing strong floor or ceiling effects) was estimated. Conditions were selected with the aim to identify one listener-specific FD per nominal *F*0 covering the entire range of expected performance across the conditions. On the lower end of the performance range, the no-SIPI and full-rate-SIPI conditions with a modulation depth of 0.1 were chosen. On the upper end, the full-rate-SIPI condition with a modulation depth of 0.7 was chosen. For the low nominal *F*0, FDs of 6%, 20%, and 46% were included and, for the high nominal *F*0, FDs of 13%, 29%, and 57% were included. The 2000-pps columns of Table II show the listener-specific FDs identified with the FD-estimation procedure for experiment 1. These columns contain the final setup of the FDs for the pitch-discrimination experiment. Note that the 1000-pps columns contain details for experiment 2 (discussed later in Sec. III A).

TABLE II. Listener-specific setup of experiment 1 (2000-pps carrier rate, pitch-discrimination experiment) and 2 (1000-pps carrier rate). FD shown in %. Lower *F*0 (L), higher *F*0 (H), and their geometric average (Avg) shown in Hz. A hyphen indicates that a listener did not show any sensitivity in that condition. Note that the smallest realizable FDs for the low nominal *F*0 differed between carrier rates (6% for 2000 pps vs 13% for 1000 pps).

	2000-pps carrier rate								1000-pps carrier rate			
	Low nominal *F*0				High nominal *F*0				Low nominal *F*0			
Listener	FD	L	H	Avg	FD	L	H	Avg	FD	L	H	Avg
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

As a consequence of this test, in the pitch-discrimination experiment, CI21 was not tested for the low nominal *F*0 and CI74 was not tested for the high nominal *F*0, because both lacked sufficient sensitivity at either FD in these conditions.

For the low nominal *F*0, the average FD (across listeners who qualified for the pitch-discrimination experiment) was 26% [standard deviation (SD) = 2.44]. For the high nominal *F*0, the average FD was 28% (SD = 1.69). The deviation of the actual average *F*0s from the targeted nominal F0s was rather small (see columns "Avg" in Table II).

6. Pitch-discrimination experiment

In the pitch-discrimination experiment, the procedure was identical to that for the FD estimation (see Sec. II A 5). Pitch discrimination was measured for all combinations of modulation depth (0.1, 0.3, 0.5, 0.7, and 0.9) and stimulus (no-SIPI, full-rate-SIPI, and half-rate-SIPI condition) for both nominal *F*0s. Additionally, pitch discrimination was tested for the loudness-balanced low-rate conditions with nominal pulse rates of 62.5, 125, and 250 pps, corresponding to the included nominal *F*0s and SIPI rates. For 62.5 and 125 pps, the FD corresponding to the low nominal *F*0 was used. For 250 pps, the FD corresponding to the high nominal *F*0 was used.

For CI17, CI18, and CI24, high-rate and low-rate conditions were tested together for both nominal *F*0s. Each high-rate condition was run 100 times and each low-rate condition was run 200 times, all presented in randomized order (3600 trials). Low-rate conditions were tested with a larger number of repetitions to account for the lower number of low-rate than high-rate conditions. Listeners were allowed to take breaks at any time. Hence, the number of blocks depended on the listener. For CI77, the two nominal *F*0s were tested on separate visits due to time restrictions. The low-rate conditions were assigned to the block with the same FD [i.e., the 63 and 125-pps condition to the block for the low nominal *F*0 (1900 trials) and the 250-pps condition to the block for the high nominal *F*0 (1700 trials)]. CI74 completed the list for the low nominal *F*0 only (1900 trials) and CI21 completed the list for the high nominal *F*0 only (1700 trials).

The results were analyzed with repeated-measures analyses of variance (ANOVAs). If necessary, the degrees of freedom were adjusted according to Greenhouse-Geisser. Two-sided Tukey's honestly significant difference (Tukey-HSD) multiple-comparison *t*-tests were used *post hoc*. The significance criterion was 5%.

B. Results

1. Overview

Figure 2 shows the pitch-discrimination *d'* scores for experiment 1 as a function of the modulation depth. Figures 2(A) and 2(B) show the averaged data with normalized standard errors (Cousineau, 2005; Morey, 2008), separated by

JASA
https://doi.org/10.1121/10.0000610

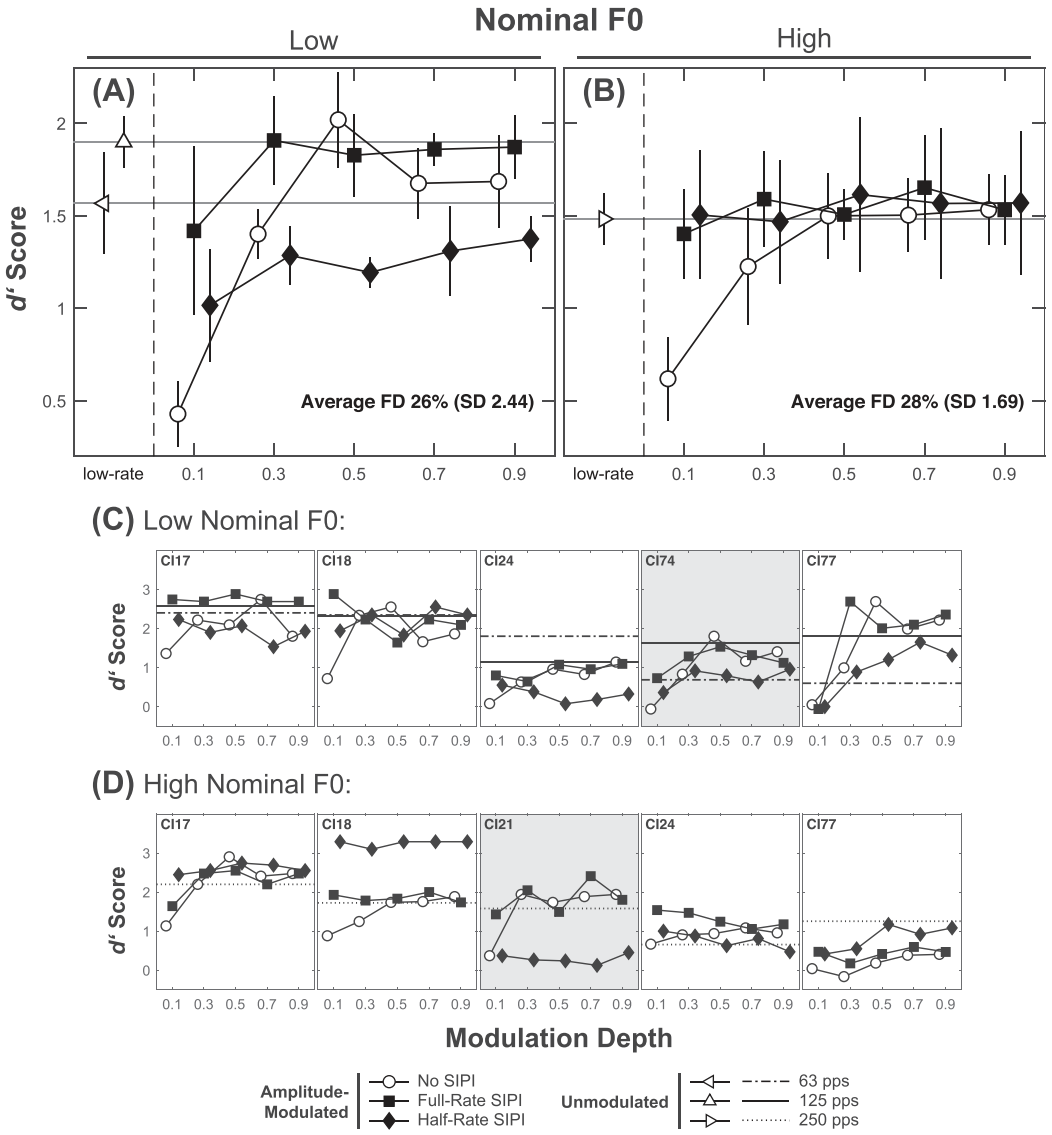


FIG. 2. Results of experiment 1. Pitch-discrimination *d'* scores with a 2000-pps carrier rate. The *d'* scores are shown for the low-rate conditions (empty triangles, shown on the abscissa at "low-rate") and as functions of the modulation depth for the no-SIPI (empty circles), the full-rate-SIPI (filled squares), and the half-rate-SIPI (filled diamonds) conditions. FD, frequency difference; SD, standard deviation. Symbols distinguish stimulus types. Top row [Panels (A) and (B)]: Group averages in black, error bars are normalized (repeated-measures) ±1 standard errors (Cousineau, 2005; Morey, 2008). Middle and bottom rows [Panels (C) and (D)]: Data from individual CI listeners; gray shading indicates that the corresponding CI listener was tested only for one nominal *F*0. Panels (A) and (C): low nominal *F*0 (around 125 Hz). Panels (B) and (D): high nominal *F*0 (around 250 Hz).

nominal *F*0 [low in Fig. 2(A) and high in Fig. 2(B)]. Figures 2(C) and 2(D) show the corresponding individual data. The different average FDs depending on the combination of carrier rate and nominal *F*0 (cf. Table II) are included in Figs. 2(A) and 2(B).

The overall pattern of results appeared to be consistent with the main hypothesis of enhanced performance for the full-rate-SIPI condition (filled squares) compared to the no-SIPI condition (empty circles), particularly at low modulation depths where performance in the no-SIPI condition was reduced. The no-SIPI and full-rate-SIPI conditions tended to yield poorer performance at the high *F*0 than at the low *F*0, despite the FD being generally larger at higher *F*0s. Half-rate-SIPI (filled diamonds) performance was broadly similar across *F*0s. However, it tended to be worse than no-SIPI and full-rate-SIPI performance at the low *F*0.

2. Effects of modulation depth and stimulus type

a. Low fundamental frequency. Our main hypothesis was that performance for the full-rate-SIPI condition should improve compared to the no-SIPI condition, particularly at

https://doi.org/10.1121/10.0000610 JASA

low modulation depths. For the half-rate-SIPI condition, the performance was expected to be similar to the full-rate-SIPI condition because here the *F*0 was assumed to be sufficiently below a rate limitation. Further, it was hypothesized that the full-rate-SIPI condition at low modulation depths would restore low-rate-like discrimination performance.

Visual inspection [Fig. 2(A)] of the data for the no-SIPI condition revealed that overall performance improved with increasing modulation depth, starting close to chance level and approaching the low-rate performance at a depth of approximately 0.5. The full-rate-SIPI condition showed roughly constant performance for modulation depths $\geq$0.3, generally matching performance for the low-rate condition. For a modulation depth of 0.1, full-rate-SIPI performance tended to be worse than for larger depths, but note the larger variability in this condition. When compared to the no-SIPI condition, the full-rate-SIPI condition showed enhanced performance at the lowest modulation depths and no systematic differences at higher modulation depths. Performance for the half-rate-SIPI condition was generally lower than for the full-rate-SIPI condition.

To test the hypotheses regarding the effects of added SIPI pulses, a two-way ANOVA was conducted on the *d'* scores including the main effect of modulation depth (0.1, 0.3, 0.5, 0.7, and 0.9) and stimulus type (no-SIPI, full-rate-SIPI, and half-rate-SIPI condition) as well as their interaction. Both main effects, modulation depth [$F(4,16) = 3.85$, $p = 0.022$, $\eta_p^2 = 0.491$] and stimulus type [$F(2,8) = 7.64$, $p = 0.014$, $\eta_p^2 = 0.657$], as well as their interaction [$F(8,32) = 2.71$, $p = 0.021$, $\eta_p^2 = 0.404$] were significant. *Post hoc* tests showed that, for the no-SIPI condition, performance was significantly worse at a modulation depth of 0.1 compared to all other depths (all $p < 0.05$). For both the full-rate-SIPI and the half-rate-SIPI condition, performance did not differ significantly across modulation depths. There were no significant performance differences between different stimulus types when comparing performance at specific modulation depths. Still, full-rate-SIPI performance was generally better than half-rate-SIPI performance ($p = 0.045$), but neither no-SIPI and full-rate-SIPI performance ($p = 0.058$) nor no-SIPI and half-rate-SIPI performance ($p = 0.489$) differed significantly.

Because no-SIPI performance improved monotonically with increasing modulation depth and because full-rate-SIPI performance was similar across modulation depths, the analysis was restricted to the lowest modulation depth of 0.1. To this end, a one-way ANOVA was conducted on the *d'* scores including the factor stimulus type with the following three levels: 125-pps low-rate, no-SIPI, and full-rate SIPI. The effect of stimulus type was significant [$F(2,8) = 9.76$, $p = 0.007$, $\eta_p^2 = 0.709$]. *Post hoc* tests showed that no-SIPI performance was worse than low-rate performance ($p = 0.001$), but neither low-rate and full-rate-SIPI performance ($p = 0.552$) nor no-SIPI and full-rate-SIPI performance ($p = 0.121$) differed significantly. Hence, these results provide only indirect evidence that full-rate SIPI pulses restored low-rate performance at low modulation depths.

In summary, as hypothesized, the analysis showed that (1) no-SIPI performance degraded for a modulation depth of 0.1 relative to higher depths, (2) performance for both the full-rate and the half-rate-SIPI conditions was similar across modulation depths, and (3) full-rate SIPI pulses restored low-rate-like performance at the lowest modulation depth. However, contrary to the hypothesis, half-rate-SIPI performance was worse than full-rate-SIPI performance. In experiment 2 (see Sec. III B), this analysis will be repeated with more data and, thus, a higher statistical power.

b. High fundamental frequency. Similar to the low-*F*0 conditions, it was hypothesized that performance in the full-rate-SIPI condition should be improved compared to the performance in the no-SIPI condition, especially at low modulation depths. Further, it was hypothesized that low-rate performance should be restored in the full-rate-SIPI condition as compared to the no-SIPI condition, particularly at low modulation depths. The half-rate-SIPI condition should circumvent a rate limitation (if present), manifesting in better half-rate-SIPI than full-rate-SIPI performance. Note that the latter hypothesis did not apply to the low-F0 conditions for which no rate-limitation circumvention was hypothesized.

Visual inspection [Fig. 2(B)] of the no-SIPI condition showed that overall performance improved with increasing modulation depth up to a depth of 0.5, approaching low-rate performance at modulation depths $\geq$0.5. The full-rate-SIPI performance was similar across modulation depths and generally matched performance for the low-rate condition. When compared to the no-SIPI condition, performance in the full-rate-SIPI condition improved at low modulation depths but did not differ systematically at higher modulation depths. Performance in the half-rate-SIPI condition was, on *average*, very similar to the full-rate-SIPI condition. However, note the larger error bars compared to the other high-*F*0 conditions indicating more heterogeneous performance across listeners [see Fig. 2(D) for the individual data].

To test the hypotheses, a two-way ANOVA was conducted on the *d'* scores including a main effect of modulation depth (0.1, 0.3, 0.5, 0.7, and 0.9) and stimulus type (no-SIPI, full-rate-SIPI, and half-rate-SIPI condition) as well as their interaction. The effect of modulation depth [$F(4,16) = 3.11$, $p = 0.045$, $\eta_p^2 = 0.437$] was significant and the effect of stimulus type [$F(1.02,4.09) = 0.28$, $p = 0.628$, $\eta_p^2 = 0.070$] was not significant. Yet, the interaction [$F(8,32) = 4.13$, $p = 0.002$, $\eta_p^2 = 0.509$] was significant. *Post hoc* tests showed that, compared to the no-SIPI condition, performance in the full-rate-SIPI condition was significantly better at modulation depths of 0.1 ($p = 0.009$) and 0.3 ($p = 0.028$). Further, at any modulation depth, performance in the half-rate-SIPI condition did not differ significantly from both the no-SIPI condition (all $p > 0.223$) and the full-rate-SIPI condition (all $p > 0.969$).

No-SIPI performance improved monotonically with increasing modulation depth and full-rate-SIPI performance was similar across modulation depths; hence, the analysis

JASA
https://doi.org/10.1121/10.0000610

was restricted to the lowest modulation depth of 0.1. To this end, a one-way ANOVA was conducted on the *d'* scores including the factor stimulus type with three levels: 250-pps low-rate, no-SIPI, and full-rate SIPI. The effect of stimulus type was significant [$F(2,8) = 8.70$, $p = 0.010$, $\eta_p^2 = 0.685$]. *Post hoc* tests showed that no-SIPI performance was worse than both full-rate-SIPI performance ($p = 0.009$) and low-rate performance ($p = 0.042$), whereas low-rate and full-rate-SIPI performance did not differ ($p = 0.960$). Thus, compared to the no-SIPI condition, full-rate SIPI pulses seemed to have restored low-rate performance at low modulation depths.

In summary, as hypothesized (hypothesis #1), the analysis showed that performance in the full-rate-SIPI condition was significantly better than in the no-SIPI condition at low modulation depths, at least at the high *F*0, but see also Sec. III B for a cross-experiment analysis, and full-rate SIPI pulses restored low-rate performance at low modulation depths. Contrary to hypothesis #2, half-rate-SIPI performance did not significantly differ from that in the no-SIPI or full-rate-SIPI conditions. However, this analysis did not reveal to what extent no-SIPI and full-rate-SIPI performance were affected by a rate limitation and if presenting the pitch cue at a lower rate (i.e., half-rate SIPI) circumvented a limitation. This is addressed in Sec. II B 3.

3. The effect of the rate

After having investigated the effects of modulation depth and stimulus type on pitch-discrimination performance, the effect of the nominal *F*0 was examined for the high-rate stimuli (no-SIPI, full-rate-SIPI, and half-rate-SIPI conditions) and the effect of the pulse rate was examined for the low-rate stimuli (63, 125, and 250 pps). To this end, the data from the four listeners (CI17, CI18, CI24, and CI77, cf. 2000-pps columns in Table II) who were tested for both nominal *F*0s were used.

Our main hypothesis was that performance in both the no-SIPI and the full-rate-SIPI conditions would degrade with increasing nominal *F*0 (and indicate a rate limitation) whereas performance in the half-rate-SIPI condition would be best among the three stimulus types for the high nominal *F*0 (and indicate a circumvention of the rate limitation). In statistical terms, an interaction of the factors nominal *F*0 and stimulus type was hypothesized. Because the rate-limitation hypothesis for no-SIPI and full-rate-SIPI stimuli was motivated by results obtained with low-rate stimuli (e.g., Ihlefeld *et al.*, 2015), a rate limitation was also hypothesized for the 250-pps *re* 125-pps low-rate conditions.

Figure 3 shows the effect of the rate for certain high-rate as well as low-rate stimuli (data are re-plotted from Fig. 2). High-rate conditions are plotted as a function of nominal *F*0 (in Hz) and low-rate conditions are plotted as a function of pulse rate (in pps). To exclude the interaction of low modulation depth and stimulus condition (see Sec. II B 2), for the high-rate stimuli only the data for the highest modulation depth of 0.9 are included.

The data in Fig. 3 showing the effect of the rate for selected high-rate stimuli (see the black symbols connected with solid lines) are in line with hypothesis #2. Performance in the no-SIPI condition (black empty circles) and in the full-rate-SIPI condition (black filled squares) degraded when the rate (i.e., nominal *F*0) was increased from 125 to 250 Hz. Conversely, performance in the half-rate-SIPI condition (black filled diamonds) improved when the rate was increased. For this subset of CI listeners, performance in the half-rate-SIPI condition tended to be worse than in the other two conditions for the low nominal

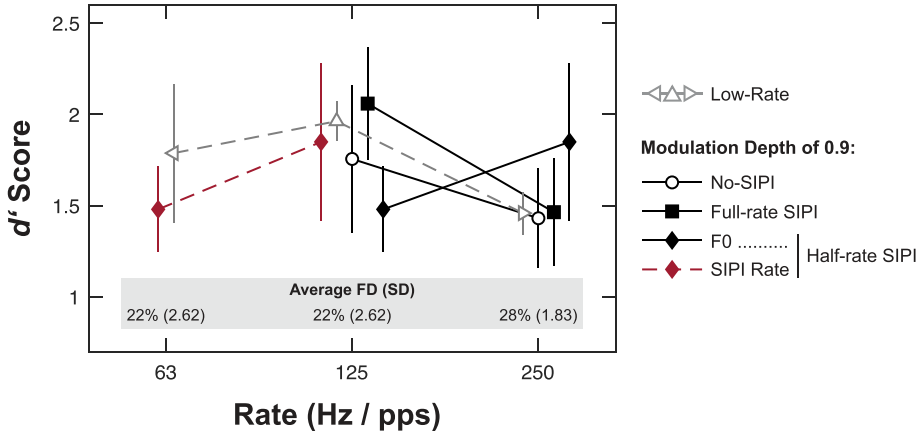


FIG. 3. (Color online) Pitch-discrimination *d'* scores for selected conditions of experiment 1 as a function of rate, i.e., *F*0 (nominally 125 and 250 Hz) for the high-rate amplitude-modulated stimuli [no-SIPI (circles), full-rate-SIPI (squares), and half-rate-SIPI (diamonds) condition; in black), pulse rate for the low-rate stimuli (nominally 63, 125, and 250 pps; gray triangles), and SIPI rate (nominally 63 and 125 pps; red diamonds) for the half-rate-SIPI condition. Filled symbols show conditions with SIPI pulses, empty symbols show conditions without SIPI pulses. Error bars show normalized ±1 standard errors. The underlying data are taken from Fig. 2 for the four-listener subset that was tested for both nominal *F*0s [panels without gray shading in Figs. 2(C) and 2(D)]. Average frequency differences (FDs, SDs in parentheses) are denoted for each rate. For the amplitude-modulated stimuli, only the data for a modulation depth of 0.9 were included. The underlying data of symbols connected with solid lines were statistically analyzed in Sec. II B 3, the data of the symbols connected with dashed lines were only visually analyzed. The same half-rate-SIPI data are plotted twice, once as a function of *F*0 and once as a function of SIPI rate. Note that, for latter case (red diamonds), the average FD denoted for 125 Hz does not match here (the 250-Hz FD is the correct one).

https://doi.org/10.1121/10.0000610 JASA

F0 but better than the other two conditions for the high nominal $F0$.[1]

To statistically test the hypothesis that half-rate SIPI pulses circumvented a rate limitation and to check for any unexpected interactions between nominal $F0$ [low (125 Hz) and high (250 Hz)], modulation depth (0.1, 0.3, 0.5, 0.7, and 0.9), and stimulus type (no-SIPI, full-rate-SIPI, and half-rate-SIPI condition), a three-way ANOVA was conducted on the d' scores including these factors and all two-way interactions and the three-way interaction. The interaction between nominal $F0$ and stimulus type was significant [$F(2,6) = 8.23$, $p = 0.019$, $\eta_p^2 = 0.733$]. Further, the interaction between modulation depth and stimulus type [$F(8,24) = 3.03$, $p = 0.008$, $\eta_p^2 = 0.541$] and the main effect of modulation depth were significant [$F(4,12) = 3.92$, $p = 0.029$, $\eta_p^2 = 0.567$], but no other main effects or interactions reached significance. The significant interaction of nominal $F0$ and stimulus type seems to support a weak version of hypothesis #3 (see Fig. 3). However, despite this interaction, *post hoc* tests did not reveal a significant advantage of the half-rate SIPI condition over the full-rate-SIPI and/or the no-SIPI condition at the high nominal $F0$, thus, not supporting a strict version of hypothesis #2. Furthermore, the significant interaction of modulation depth and stimulus type found here for a subset of four CI listeners was also found in the analyses above conducted separately for each nominal $F0$s and with five CI listeners (see also the extended analysis in Sec. III B).

In addition to the high-rate data, Fig. 3 also includes results for the low-rate stimuli (gray empty triangles). The low-rate data show that pitch-discrimination performance was, in fact, a *non-monotonic* function of rate, with the best performance at the intermediate rate (125 pps) and worse performance both at lower and higher rates (63 and 250 pps, respectively). Degrading low-rate performance with an increasing pulse rate from 125 to 250 pps is consistent with the no-SIPI and the full-rate-SIPI conditions. Improving low-rate performance with increasing pulse rate from 63 to 125 pps supports the idea that the lower half-rate SIPI performance reflects a perceived pitch at half the $F0$ because the SIPI rate rather than the $F0$ dominated the percept. In particular, this inverse effect for the half-rate-SIPI condition may be understood by simply re-plotting the data as a function of SIPI rate rather than $F0$ (red filled diamonds). However, note the slightly different FDs used for the low-rate condition compared to the half-rate-SIPI condition at 125 pps.

In summary, there was a significantly different effect of the nominal $F0$ for the half-rate-SIPI condition compared to all other tested low-rate and high-rate conditions. Although the half-rate SIPI condition was less affected by a rate limitation (for the subset of CI listeners, it actually showed *higher* performance at 250 than at 125 Hz), it did not outperform the no-SIPI and full-rate SIPI conditions at 250 Hz. Between 63 and 250 pps, pitch-discrimination performance was non-monotonic with the best performance at 125 pps. This indirectly suggests that for the half-rate-SIPI condition, the SIPI rate may have dominated the pitch percept, thus lowering the perceived pitch. However, a more direct measurement is necessary to support this conclusion.

III. EXPERIMENT II: 1000-PPS CARRIER RATE

In multi-electrode configurations necessary for speech understanding and generally employed in clinical applications, the per-electrode carrier rates are usually closer to 1000 pps. Per-channel rates as high as 2000 pps, as used in experiment 1, are rarely used (Wouters *et al.*, 2015). However, it is unclear which single-electrode rate would best compare to multi-electrode stimulation used with clinically applied stimulation strategies. Such strategies potentially drive any group of neurons with much higher composite stimulation rates than corresponding to the channel-specific rates, due to the overlap of electric fields from neighbouring electrodes (e.g., Boulet *et al.*, 2016).

To determine whether the effects of modulation depth and of the different SIPI conditions generalized to a pulse rate more typical of the per-channel pulse rates used clinically, experiment 2 used a carrier rate of 1000 pps, but tested only the low $F0$ (nominally 125 Hz). Otherwise, the experimental setup was identical to experiment 1. It was hypothesized that the carrier rate would not influence pitch-discrimination sensitivity (Green *et al.*, 2012), as long as all occurring $F0$s are sufficiently oversampled (e.g., Wilson *et al.*, 1997) and as long as the carrier rate is well above the rate limitation for temporal pitch. Furthermore, data collected at 1000 pps would allow for a better comparison with the results for ITD sensitivity in Srinivasan *et al.* (2018, 2020), which were collected at 1000 pps.

The data from this experiment also allowed an extension of the analysis of the effects of modulation depth and stimulus type from experiment 1 (see Sec. II B 2 a) with higher power due to the doubled amount of data.

A. Methods

All listeners tested with the low nominal $F0$ in experiment 1 (all but CI21) also participated in experiment 2. Only high-rate stimuli [no-SIPI, full-rate-SIPI, and half-rate-SIPI conditions, cf. Figs. 1(A)–1(C)] were used. To ensure six-times oversampling of the temporal envelope as in experiment 1, only the low nominal $F0$ of 125 Hz was tested. The SIPI fraction was set to 6%, which resulted in the same time gap of 60 μs between the regular and the SIPI pulse as for the 2000-pps carrier rate in experiment 1. Note that the smallest realizable FD was 13% for the 1000-pps carrier, in contrast to 6% for the 2000-pps carrier. The individualized FDs for experiment 2 were based on those for experiment 1 and are shown in the rightmost column of Table II. The average FD was 25% (SD = 1.83). In all other aspects, stimuli, apparatus, and tasks were identical to the low nominal $F0$ tested in experiment 1.

The 1000-pps stimuli were loudness-balanced to the same (2000-pps) reference used in experiment 1 by a two-stage approach. The 1000-pps, 125 Hz-$F0$, 0.5-modulation-depth stimulus without SIPI pulses (no-SIPI condition) was formally loudness-balanced. The amplitudes for the other 1000-pps conditions were informally loudness-balanced. To this end, the 1000-pps amplitudes were extrapolated.

JASA
https://doi.org/10.1121/10.0000610

First, the difference in loudness between the 2000-pps reference and a 1000-pps reference ($F0$ 125 Hz, modulation depth 0.5, no-SIPI condition) was formally measured using the staircase procedure described in Sec. II A 4. Second, this difference was added to the loudness-balanced amplitudes of the 2000-pps targets from experiment 1 to obtain the loudness-balanced amplitudes of the 1000-pps targets. To validate this extrapolation, listeners were repeatedly presented with selected conditions and informally asked to indicate if the 2000-pps reference and 1000-pps target were equally loud. If necessary, the extrapolated amplitudes were adjusted.

B. Results

It was hypothesized that the pattern of results observed in experiment 1 would generalize to a 1000-pps carrier rate. Consequently, the effects of modulation depth and stimulus type were expected to be similar as found for the 2000-pps carrier rate and the low nominal $F0$ in experiment 1. In particular, it was hypothesized that performance in the full-rate-SIPI condition would be better than in the no-SIPI condition, especially at low modulation depths. It was further hypothesized that full-rate and half-rate SIPI pulses would yield similar performance since all occurring $F0$s and SIPI rates were sufficiently below the assumed rate limitation.

Visual inspection of the averaged results in Fig. 4(A) showed that performance in the no-SIPI condition improved with increasing modulation depth up to a depth of 0.5 while saturating for higher depths. Performance in the full-rate-SIPI and the half-rate-SIPI conditions seemed to be roughly constant across modulation depths. In line with the hypothesis, full-rate-SIPI performance seemed to outperform no-SIPI performance at low modulation depths. Half-rate-SIPI performance was generally worse than full-rate-SIPI performance and tended to be similar to no-SIPI performance at low modulation depths while being worse at higher depths. Visual comparison of the results with those of the low nominal $F0$ of experiment 1 at 2000-pps carrier rate [Fig. 2(A)] revealed broadly similar performance across carrier rates.

To test the hypotheses, a three-way ANOVA was conducted on the d' scores including the data collected in both experiments for the low nominal $F0$. A main effect of modulation depth (0.1, 0.3, 0.5, 0.7, and 0.9), stimulus type (no-SIPI, full-rate-SIPI, and half-rate-SIPI condition), and carrier rate (1000 and 2000 pps) was included, as well as all two-way interactions and the three-way interaction. Note that different FDs were used for the two carrier rates, which could have been a possible confounding factor, despite the fact that, averaged across listeners, the difference in FD was quite small. However, the main interest was in whether there would be an interaction between carrier rate and the effects of modulation depth and stimulus type. No such interactions were observed [carrier rate × modulation depth: $F(4,16) = 1.79$, $p = 0.180$, $\eta_p^2 = 0.309$; carrier rate × stimulus type: $F(2,8) = 1.23$, $p = 0.340$, $\eta_p^2 = 0.236$; carrier rate × modulation depth × stimulus type: $F(8,32) = 0.75$, $p = 0.645$, $\eta_p^2 = 0.159$], and,

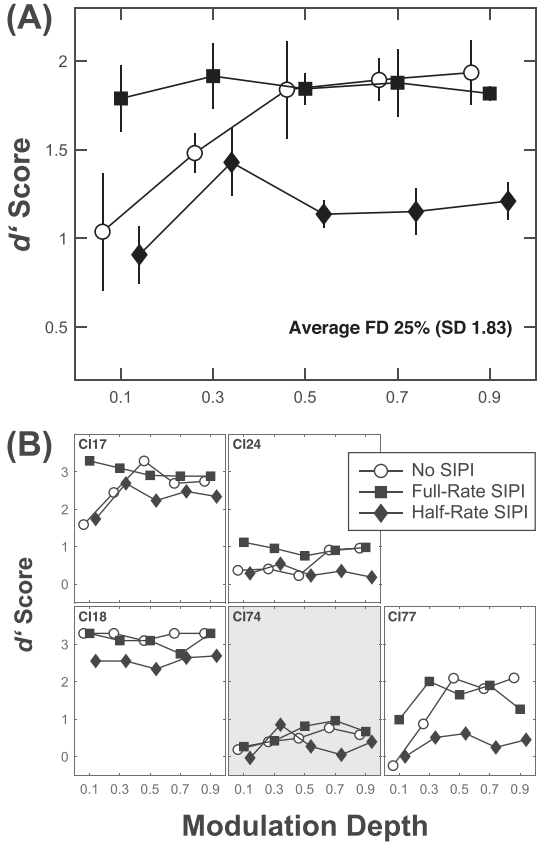


FIG. 4. Results of experiment 2. Pitch-discrimination d' scores with a 1000-pps carrier rate and for the low nominal $F0$ (around 125 Hz) only. The d' scores are shown as functions of the modulation depth for the no-SIPI (empty circles), the full-rate-SIPI (filled squares), and the half-rate-SIPI (filled diamonds) conditions. FD, frequency difference; SD, standard deviation. Panel (A): Group averages in black, error bars are normalized (repeated-measures) ±1 standard errors. Panel (B): Data from individual CI listeners; gray shading indicates that the corresponding CI listener was tested only for the low nominal $F0$ in experiment 1.

in fact, the main effect of carrier rate was not significant either [$F(1,4) = 0.06$, $p = 0.813$, $\eta_p^2 = 0.016$]. Hence, the analysis did not reveal any evidence of the carrier rate affecting pitch-discrimination performance, regardless of whether SIPI pulses were inserted or not. Still, the effects of modulation depth [$F(4,16) = 3.89$, $p = 0.021$, $\eta_p^2 = 0.494$] and stimulus type [$F(2,8) = 15.71$, $p = 0.002$, $\eta_p^2 = 0.797$] were significant, but the interaction of modulation depth and stimulus type was also significant [$F(8,32) = 4.36$, $p = 0.001$, $\eta_p^2 = 0.522$]. *Post hoc* tests showed that, at a modulation depth of 0.1, performance in the full-rate-SIPI condition was significantly better than in both the no-SIPI ($p = 0.026$) and the half-rate-SIPI condition ($p = 0.016$). Further, performance in the full-rate-SIPI condition was generally better than in the half-rate-SIPI condition ($p = 0.020$).

At a modulation depth of 0.1, although not reflected in the statistical results, CI18's discrimination performance in the no-SIPI condition differed considerably between carrier rates [Fig. 2(C) vs Fig. 4(B)]. With the 1000-pps carrier,

https://doi.org/10.1121/10.0000610 JASA

performance was perfect already at the lowest modulation depth which might suggest an effect of the carrier rate. However, a different explanation was considered to be more reasonable. As denoted in Table II, CI18 was tested with an FD of 6% for the 2000-pps carrier whereas an FD of 13% had to be used for the 1000-pps carrier due to limitations in the stimulus generation (see Sec. III A). Hence, the FD was more than doubled in experiment 2 and the task was too easy for the 1000-pps carrier.

In summary, experiment 2 showed very similar effects of modulation depth and stimulus type compared to the low-*F*0 data in experiment 1, thus no indication of an effect of the carrier rate. When pooling the data across carrier rates, the full-rate-SIPI condition outperformed the no-SIPI condition at a modulation depth of 0.1 and the half-rate-SIPI condition across all depths. Thus, the ITD data of Srinivasan *et al.* (2018, 2020), collected at a carrier rate of 1000 pps, can be equally compared to the present experiment collected at the same rate or to experiment 1 collected at 2000 pps.

IV. OVERALL DISCUSSION

Temporal-envelope pitch-discrimination performance was measured for CI listeners stimulated at a single electrode. In experiment 1, the stimuli were 2000-pps amplitude-modulated pulse trains (a) without SIPI pulses (no-SIPI condition), as conveyed by current CI stimulation paradigms, (b) with SIPI pulses inserted at the full envelope rate (full-rate-SIPI condition), or (c) with SIPI pulses inserted at half the envelope rate (half-rate-SIPI condition). A wide range of modulation depths (from 0.1 to 0.9) was tested. *F*0s corresponded to male (nominally 125 Hz) and female (nominally 250 Hz) talkers. In addition, conditions with low-rate unmodulated pulse trains with rates corresponding to all combinations of *F*0s and SIPI rates were tested. A follow-up experiment (experiment 2) tested the low-*F*0 conditions at a 1000-pps carrier rate closer to clinically used carrier rates.

A. The effects of modulation depth and stimulus type

1. Amplitude modulation only

It was hypothesized that pitch-discrimination performance with amplitude-modulated stimuli without SIPI pulses (no-SIPI condition) would be weak at low modulation depths. Indeed, the performance was close to chance level at a modulation depth of 0.1 and approached high scores ($d' \geq 1$) at a modulation depth larger than 0.3. The low performance at low modulation depths is reasonable, assuming that precise entrainment of neural responses to the stimulus' envelope requires a minimum onset slope and off-time within each modulation period (e.g., Hancock *et al.*, 2017).

Geurts and Wouters (2001) investigated single-electrode pitch-discrimination sensitivity for sinusoidally amplitude-modulated pulse trains and found a saturation of discrimination performance at modulation depths larger than 0.2. This is in line with the present results, suggesting that the particular envelope shape (i.e., a full-wave rectified sinusoid) used in this study yielded similar pitch sensitivity, and consistent with other studies showing that, in electric hearing, the envelope shape has little effect on amplitude-modulation rate discrimination (Landsberger, 2008; Kreft *et al.*, 2010; Hancock *et al.*, 2017).

2. SIPI pulses inserted at full rate

It was hypothesized that inserting SIPI pulses would enhance pitch-discrimination performance, particularly at low modulation depths where temporal-envelope pitch cues are weak. Indeed, compared to the no-SIPI condition, performance improved in the full-rate-SIPI condition for modulation depths of 0.1 (both nominal *F*0s) and 0.3 (high nominal *F*0 only, see Secs. II B 2 and III B). The transient and periodical timing cues evoked by the SIPI pulses in the full-rate-SIPI condition, therefore, have the potential to enhance both female and male-*F*0 cues particularly at low modulation depths.

This improvement is of particular interest when considering the modulation depths usually found in electric hearing. To this end, envelope statistics of sentences from both English and German speech corpora were calculated.[2] Results showed that clinical CI processors typically produce modulation depths ranging between 0.1 and 0.4 with an average of 0.25. Note that in everyday situations, these modulation depths may be even lower, for example, due to noise interference or reverberation (e.g., Sayles and Winter, 2008). Assuming that the periodicity of the signal can still be robustly detected by a signal-processing algorithm, application of full-rate SIPI pulses in a CI stimulation paradigm seems to have the potential to provide salient and gender-sensitive *F0* cues even under adverse listening conditions.

The results are broadly consistent with a recent study on the effect of the SIPI approach on ITD sensitivity with unmodulated high-rate pulse trains by Srinivasan *et al.* (2018). In an attempt to characterize the mechanism underlying the SIPI effect on ITD sensitivity, Srinivasan *et al.* compared performance between an unmodulated pulse train with SIPI pulses and (1) a so-called *attenuated* pulse train where the short-term power of SIPI-pulse pairs was reduced to equal that of the regular (single) carrier pulses, and (2) a so-called *enhanced* pulse train where the short-term power of periodically selected regular carrier pulses in a periodic pulse train was enhanced to equal the power of SIPI-pulse pairs. Their results showed comparable performance for the SIPI and the *enhanced* condition, while the *attenuated* condition revealed chance-level performance, similar to a high-rate pulse train without SIPI pulses. Srinivasan *et al.* concluded that the insertion of SIPI pulses introduced a form of artificial amplitude modulation in the neural activation pattern which in turn triggered temporally precise neural spikes. It is reasonable to assume the same modulation enhancement mechanism could explain the SIPI effect on temporal-pitch sensitivity found in the present study. The constant SIPI performance across modulation depths seems to suggest that SIPI pulses trigger similar spike patterns

JASA
https://doi.org/10.1121/10.0000610

regardless of the modulation depth of the amplitude-modulated signal.

When assuming that SIPI pulses enhance temporal-pitch cues by means of modulation enhancement, it appears interesting to compare the full-rate-SIPI condition to signal-processing paradigms that were explicitly designed to enhance temporal-envelope pitch cues by modifying the envelope shape. Unfortunately, modifications of the envelope shape potentially affect speech perception, and the available paradigms actually paint a mixed picture. Some paradigms enhanced pitch sensitivity but degraded transmission of speech information (Green *et al.*, 2004; Green *et al.*, 2005), whereas other paradigms improved performance in pitch-related tasks but had (almost) no effect on speech recognition (Vandali *et al.*, 2005; Laneau *et al.*, 2006; Milczynski *et al.*, 2009; Milczynski *et al.*, 2012; Francart *et al.*, 2015; Vandali and van Hoesel, 2011, 2012). Envelope-sharpening approaches have also been shown to enhance ITD sensitivity (e.g., Laback *et al.*, 2011; Monaghan and Seeber, 2016). However, as for the SIPI approach, the effects of those approaches on CI listeners' speech understanding have not been studied sufficiently so far (see also Sec. IV E).

The results from Srinivasan *et al.* (2018) at first hand suggest that the SIPI approach is "just" some form of modulation enhancement (their *enhanced* condition). Further, it may even appear to be less flexible than a general enhancement approach because, when considering a multi-electrode stimulation paradigm such as CIS, a SIPI pulse needs to be paired with the regular carrier pulse, occupying a longer time slot in the tightly interleaved multi-electrode pulse patterns. However, a potential disadvantage of the *enhanced* approach may be that the increase of the signal's peak factor reduces the maximally allowed level of the non-enhanced stimulus pulses as given by the maximum comfort level. Further, Srinivasan *et al.* (2018) showed that the *enhanced* approach differs from the SIPI approach in one important aspect, that is, loudness. In fact, while providing similar ITD sensitivity, pulse *enhancement* elicited more loudness increase than the SIPI approach. Along this line, Buechel *et al.* (2018) hypothesized that, for ITD perception, the subset of neurons showing a SIPI effect might have differed from the subset of neurons that contributed to loudness. Because Srinivasan *et al.* (2018) studied the loudness effect with unmodulated pulse trains only, a study with amplitude-modulated pulse trains would be required the determine the extent of the difference in loudness perception between the two approaches.

3. SIPI pulses inserted at half rate

For low $F0$s (not affected by a potential rate limitation), it was hypothesized that performance in the half-rate-SIPI condition would be similar to that in the full-rate-SIPI condition whereas, for higher $F0$s (possibly affected by a rate limitation), it was hypothesized that the performance in the half-rate-SIPI condition would be better than in the full-rate-SIPI condition (thereby circumventing the rate limitation). While the effect of the $F0$ was indeed significantly different for the half-rate-SIPI condition than for the no-SIPI and half-rate SIPI conditions, the performance was not as anticipated. For the low $F0$, the half-rate-SIPI condition performed significantly worse than the full-rate-SIPI condition, and for the high $F0$, the performance was more listener-specific and, on *average*, it did not outperform the no-SIPI and full-rate-SIPI conditions. A comparison with the low-rate stimuli (see Fig. 3) provided some hints that listeners might have relied on the pitch conveyed by the SIPI pulses rather than by the $F0$-based envelope modulation. Note that performance was generally worse at 63 Hz than at 125 Hz. Despite differences in approach and methodology, this conclusion would be consistent with the pitch-matching results from Pieper and Bahmer (2019) obtained between unmodulated 400-pps pulse trains with and without SIPI pulses inserted at half the pulse rate. Finally, the dominance of the SIPI rate would be consistent with the general dominance of the lower rate as shown with unmodulated irregular-rate pulse trains (Carlyon *et al.*, 2002). Despite these inductions, a more direct estimation of the perceived pitch of the half-rate-SIPI condition is clearly required in future studies.

B. Rate limitation

For the no-SIPI, the full-rate-SIPI, and the low-rate conditions, performance was hypothesized to degrade with increasing $F0$ from 125 to 250 Hz or pps, respectively. For the half-rate-SIPI condition, assuming dominance of the SIPI cue, no such limitation was expected at the high $F0$. Indeed, such a pattern of results (see Sec. II B 3) was found with a matched subset of CI listeners.

Although the rate effect was overall weak, these results are consistent with numerous studies showing that rate pitch is affected by a pulse-rate limitation which starts between 200 and 300 pps (e.g., Townshend *et al.*, 1987; Kong *et al.*, 2009; Vandali *et al.*, 2013; Ihlefeld *et al.*, 2015). Moreover, the results for the no-SIPI condition at high modulation depths are consistent with Chatterjee and Oberzut (2011), who reported a decrease in amplitude-modulation rate-discrimination sensitivity for reference modulation rates above 100 Hz.

Apart from the low number of CI listeners, two factors were probably responsible for the small effect of the rate. First, the high $F0$ was only around 250 Hz. This is not particularly high and performance might only be slightly influenced by the rate limitation. Second, the FDs were not identical across rates, not even within participants. The generally slightly larger FDs at the high nominal $F0$ (cf. Table II) might have counteracted the effect of the rate limitation.

While no explicit hypothesis was formulated for the 63-pps low-rate condition, a trend towards lower performance than in the 125-pps low-rate condition was observed. This trend was similarly observed in the half-rate-SIPI condition. This suggests the presence of a "low-rate" limitation (see Fig. 3), which appears to be consistent with previous studies

https://doi.org/10.1121/10.0000610 JASA

reporting degrading pitch sensitivity towards low rates (at least for a subset of CI listeners) for both unmodulated low-rate pulse trains (Kong *et al.*, 2009; Stahl *et al.*, 2016) and amplitude-modulated high-rate pulse rates (Kong *et al.*, 2009; Chatterjee and Oberzut, 2011).

C. Comparison of unmodulated low-rate and amplitude-modulated high-rate pulse trains

For both nominal *F*0s, it was hypothesized that performance with full-rate SIPI pulses would restore low-rate performance. This hypothesis was confirmed.

As discussed in Sec. IV A 2, SIPI pulses introduce transient, temporally repeating timing cues and can be considered as a form of modulation enhancement. The results suggest that the SIPI-induced modulation enhancement is sufficient to allow for low-rate-like temporal-pitch sensitivity. Vandali *et al.* (2013) let CI listeners pitch-match unmodulated low-rate pulse trains to 1800-pps pulse trains with two different types of amplitude modulation, either more transient exponential-decay modulation or less transient sinusoidal amplitude modulation. They showed that amplitude-modulated stimuli yielded overall higher pitches than the unmodulated low-rate pulse trains, but, the pitch converged towards that of the low-rate pulse trains when the modulation depth was increased. The pitch with the more transient exponential-decay modulation was generally more similar to the low-rate pitch. SIPI-induced modulation enhancement can be assumed to sharpen the effective modulation. Hence, the results obtained by Vandali *et al.* in combination with the findings of the present study suggest that (1) the pitch of the real-world sounds conveyed by typical stimulation paradigms varies substantially depending on the modulation depth and (2) imposing SIPI pulses on amplitude-modulated pulse trains produces a lower and more stable pitch percept, particularly for low modulation depths and therewith better resembling the pitch elicited by low-rate stimulation.

Our finding that no-SIPI performance at high modulation depths was similar to low-rate performance is consistent with results found for ITD sensitivity where the sensitivity for a fully amplitude-modulated pulse train can approach the sensitivity for an unmodulated pulse train at the corresponding rate (e.g., van Hoesel *et al.*, 2009).

D. The effect of the carrier rate

Because all occurring amplitude-modulation rates were oversampled at least by a factor of six, an effect of the carrier rate as a result of worsened sampling of the envelope with the 1000-pps carrier (experiment 2) compared to the 2000-pps carrier (experiment 1) was expected to be unlikely. Consequently, it was hypothesized that the carrier rate would not have any effect on temporal-pitch sensitivity. In line with the hypothesis, no systematic effect of the carrier rate was found when comparing the data for the low nominal *F*0 from experiment 1 and the data of experiment 2 (see Sec. III B).

Green *et al.* (2012) measured just-noticeable differences for CI amplitude-modulation rate discrimination for carrier pulse rates ranging from 482 to 5787 pps and for modulation rates of 10 and 100 Hz (at least about five-times oversampling). They did not find any significant effect of the carrier rate as long as they expressed the modulation depth relative to the dynamic range (as done in the present study). This is consistent with the results of this study. Interestingly, both the present results and those obtained by Green *et al.* (2012) do not support the physiological results of Kirby and Middlebrooks (2010) for anesthetized guinea pigs, who showed temporal precision to degrade by an order of magnitude when increasing the pulse rate from 254 to 4069 pps.

Despite sufficient envelope sampling, refractory effects are assumed to be generally stronger for a 2000-pps carrier rate than for a 1000-pps carrier rate (e.g., Boulet *et al.*, 2016). The present results suggest that, at least for the *F*0s used in the comparison, refractory effects do not importantly affect envelope-based pitch perception at the tested carrier rates. The lack of an effect of the carrier rate may be relevant for future clinical applications because envelope-based stimulation paradigms such as CIS typically use carrier rates within the tested range.

E. Relevance for future cochlear implant stimulation paradigms

Current clinically applied envelope-based CI stimulation paradigms provide high levels of speech understanding in quiet (e.g., Wilson *et al.*, 1991). Because the acoustic temporal fine-structure is discarded, temporal-pitch sensitivity (e.g., Vongphoe and Zeng, 2005; Kalathottukaren *et al.*, 2015) and ITD sensitivity (e.g., Laback *et al.*, 2004) are impaired with clinical signal processors.

To improve encoding of temporal features, on the one hand, stimulation paradigms which provide low-rate electric fine-structure coding, at least at apical electrodes, were developed (e.g., Churchill *et al.*, 2014; fine-structure processing, FSP: Zierhofer, 2003; FSP on four electrodes, FS4: Riss *et al.*, 2014; peak-derived timing, PDT: van Hoesel and Tyler, 2003). Recently, Thakkar *et al.* (2018) showed that including low-rate stimulation at one middle or *basal* interaural electrode pair along four interaural high-rate electrode pairs provided ITD sensitivity comparable to when all electrode pairs provided low-rate stimulation. These results suggest a potential benefit of explicit temporal coding in non-apical regions with mixed-rate stimulation.

On the other hand, stimulation paradigms using low pulse rates across the entire electrode array were developed to improve temporal-pitch and ITD sensitivity (e.g., fundamental asynchronous stimulus timing, FAST: Smith *et al.*, 2014; peak-picking, PP: Williges *et al.*, 2018). Note that low carrier rates have been shown to degrade speech understanding (e.g., Loizou *et al.*, 2000; Arora *et al.*, 2009; Williges *et al.*, 2018). While a deterioration was found even if the pulse timing carried information about the speech signal (Williges *et al.*, 2018), performance with low-rate

JASA
https://doi.org/10.1121/10.0000610

strategies may depend on the individual listener and the particular speech-coding paradigm.

So far, it has been shown that inserting SIPI pulses enhances single-electrode temporal-pitch sensitivity with amplitude-modulated stimuli (present study) as well as ITD sensitivity with both unmodulated (Srinivasan *et al.*, 2018) and amplitude-modulated stimuli (Srinivasan *et al.*, 2020). Thus, developing a stimulation paradigm which uses CIS-like envelope stimulation with SIPI pulses strategically inserted seems to be a promising approach to improve speech understanding in noise by means of better segregation of target and interferers based on ITD and/or temporal-pitch cues. Future studies are required to examine the feasibility of such paradigms.

For everyday stimuli, envelope peaks have to be detected to determine the timing of SIPI pulses. Further, those envelope peaks have to be sampled by a carrier pulse to place a SIPI pulse after it. With the static approach used in this study, this will only be the case if the carrier rate is an integer multiple of the envelope repetition rate. Thus, the temporal resolution is defined by the carrier rate and would be 1000 μs for a typical stimulation strategy using a 1000-pps channel rate. It is thus unlikely that such a constant-sampling approach used in this study would provide sufficient ITD sensitivity without further modifications. To this end, the findings of Egger *et al.* (2017) on the tolerable inaccuracy in encoding ITD cues with low-rate stimulation might be considered when determining the required precision and the necessary modifications of a future stimulation paradigm because SIPI stimulation is comparable to low-rate stimulation with regard to both pitch (this study) and ITD cues (Srinivasan *et al.*, 2018, 2020).

The SIPI approach, when selectively applied to *F*0-based amplitude-modulation patterns, differs from existing envelope-sharpening approaches in several aspects. First, apart from the single extra pulse at the envelope peak, it does not further modify the envelope shape. It thus has the potential to be more selective, promising less alteration of the temporal envelope of speech. Second, by fixing the SIPI-pulse position at an envelope's peak, information augmented by the SIPI pulse is clearly linked with the acoustic information of the sound. While these conjectures have not yet been examined in behavioral experiments, such data will be required to determine to what extent the SIPI approach reduces the trade-off between temporal-pitch and ITD perception, on the one hand, and speech perception, on the other hand, in future CI stimulation paradigms.

V. SUMMARY AND CONCLUSIONS

This study investigated temporal-pitch sensitivity in CI listeners when inserting SIPI pulses into high-rate amplitude-modulated single-electrode pulse trains. The outcome can be summarized as follows:

(1) Envelope-based pitch sensitivity (no-SIPI condition) deteriorated for low modulation depths. This effect likely limits access to pitch cues in the temporal envelope with current CIs in everyday situations where the effective modulation depth is low (e.g., due to environmental noise or reverberation). At moderate to high modulation depths, envelope-based pitch sensitivity was comparable to that found with low-rate stimulation.
(2) The strategic insertion of SIPI pulses at the rate corresponding to the repetition rate of the temporal envelope (full-rate-SIPI condition) is effective in enhancing both male and female *F*0 cues, with a particular benefit at low modulation depths. The transient and periodical timing cues introduced by the SIPI pulses evoked a form of modulation enhancement, which seems to be able to restore pitch sensitivity to a level observed for unmodulated low-rate pulse trains.
(3) Performance in the no-SIPI and the full-rate-SIPI conditions tended to be lower at high *F*0s. SIPI pulses inserted at the rate corresponding to half of the envelope rate (half-rate-SIPI condition) did not circumvent this effect but resulted in high across-listener variability and worse performance at low *F*0s. Thus, there seems to be not much potential for the use of half-rate SIPI pulses.
(4) For *F*0s around 125 Hz, no effect of the carrier rate between 1000 and 2000 pps was found. This suggests that pitch-discrimination sensitivity was not affected by refractory effects for carrier inter-pulse intervals between 500 and 1000 μs.

ACKNOWLEDGMENTS

We are grateful to our CI listeners for their patience and enthusiasm while conducting the tedious tests. We thank Michael Mihocic for experimental hardware and software support as well as him and Maike Ferber for assistance with collecting parts of the data. We thank the Institute of Ion Physics and Applied Physics of the Leopold-Franzens-University, Innsbruck, Austria, for providing the equipment for direct electric stimulation. We further thank the associate editor Joshua G. Bernstein and two anonymous reviewers for their fruitful comments on earlier versions of the manuscript. This study was supported by the National Institutes of Health, grant R01 DC 005775. Additional support was provided by the OeAD, grant MULT_DR 11/2017n.

[1]Note that with the full five-listener data set [cf. Fig. 2(B)], at the high nominal *F*0, performance in the half-rate-SIPI condition was more similar to the performance in the other two conditions. This difference emerged because CI21, who performed at chance in the half-rate SIPI condition [cf. Fig. 2(D)], was not included in the four-listener subset (for details, see Sec. IV A 3). Note that CI21 was the only pre-lingually deafened listener in this experiment.

[2]English sentences were taken from the TIMIT corpus (Garofolo *et al.*, 1993), German sentences were taken from the Kiel corpus (Kohler, 1996). Both male and female speakers were included. The speech data were fed to emulations of both MED-EL and Cochlear implementations of the CIS paradigm including loudness mapping. The *F*0-modulation-depth statistic was derived from data for the MED-EL processor and two electrode channels, one apical (608 Hz) and one middle (1403 Hz). The statistics were

https://doi.org/10.1121/10.0000610

JASA

further similar to the Cochlear processor. The analysis was included in the proposal of the NIH project "Bilateral Cochlear Implants: Physiology and Psychophysics," Grant No. R01 DC 005775.

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.

Boulet, J., White, M., and Bruce, I. C. (**2016**). "Temporal considerations for stimulating spiral ganglion neurons with cochlear implants," J. Assoc. Res. Otolaryngol. **17**, 1–17.

Buechel, B. D., Hancock, K. E., Chung, Y., and Delgutte, B. (**2018**). "Improved neural coding of ITD with bilateral cochlear implants by introducing short inter-pulse intervals," J. Assoc. Res. Otolaryngol. **19**, 681–702.

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**, 2997–3008.

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**, 621–633.

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

Chen, J., Yang, H., Wu, X., and Moore, B. C. J. (**2018**). "The effect of F0 contour on the intelligibility of speech in the presence of interfering sounds for Mandarin Chinese," J. Acoust. Soc. Am. **143**, 864–877.

Churchill, T. H., Kan, A., Goupell, M. J., and Litovsky, R. Y. (**2014**). "Spatial hearing benefits demonstrated with presentation of acoustic temporal fine structure cues in bilateral cochlear implant listeners," J. Acoust. Soc. Am. **136**, 1246–1256.

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

Cutler, A., Dahan, D., and van Donselaar, W. (**1997**). "Prosody in the comprehension of spoken language: A literature review," Lang. Speech **40**, 141–201.

Darwin, C. J. (**2005**). "Pitch and auditory grouping," in *Pitch*, edited by C. J. Plack, R. R. Fay, A. J. Oxenham, and A. N. Popper (Springer, New York), pp. 278–305.

Dorman, M., and Wilson, B. (**2004**). "The design and function of cochlear implants," Am. Sci. **92**, 436–445.

Egger, K., Majdak, P., and Laback, B. (**2017**). "Binaural timing information in electric hearing at low rates: Effects of inaccurate encoding and loudness," J. Acoust. Soc. Am. **141**, 3164–3174.

Francart, T., Osses, A., and Wouters, J. (**2015**). "Speech perception with F0mod, a cochlear implant pitch coding strategy," Int. J. Audiol. **54**, 424–432.

Garofolo, J. S., Lamel, L. F., Fisher, W. M., Fiscus, J. G., Pallett, D. S., Dahlgren, N. L., and Zue, V. (**1993**). *TIMIT Acoustic Phonetic Continuous Speech Corpus* (Linguistic Data Consortium, Philadelphia, PA, USA).

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

Green, T., Faulkner, A., and Rosen, S. (**2004**). "Enhancing temporal cues to voice pitch in continuous interleaved sampling cochlear implants," J. Acoust. Soc. Am. **116**, 2298–2310.

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**, 221–230.

Green, T., Faulkner, A., Rosen, S., and Macherey, O. (**2005**). "Enhancement of temporal periodicity cues in cochlear implants: Effects on prosodic perception and vowel identification," J. Acoust. Soc. Am. **118**, 375–385.

Hancock, K. E., Chung, Y., and Delgutte, B. (**2012**). "Neural ITD coding with bilateral cochlear implants: Effect of binaurally coherent jitter," J. Neurophysiol. **108**, 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**, 771–788.

Houtsma, A. J. M. (**1995**). "Pitch Perception," in *Hearing*, edited by B. C. J. Moore (Academic Press, London, UK), pp. 267–295.

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**, 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**, 641–652.

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

Kalathottukaren, R. T., Purdy, S. C., and Ballard, E. (**2015**). "Prosody perception and musical pitch discrimination in adults using cochlear implants," Int. J. Audiol. **54**, 444–452.

Kirby, A. E., and Middlebrooks, J. C. (**2010**). "Auditory temporal acuity probed with cochlear implant stimulation and cortical recording," J. Neurophysiol. **103**, 531–542.

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

Kohler, K. J. (**1996**). "Labelled data bank of spoken standard German: The Kiel corpus of read/spontaneous speech," in *Proceeding of Fourth International Conference on Spoken Language Processing, ICSLP ';96*, October 3–6, Philadelphia, PA, pp. 1938–1941.

Kong, Y.-Y., and Carlyon, R. P. (**2010**). "Temporal pitch perception at high rates in cochlear implants," J. Acoust. Soc. Am. **127**, 3114–3123.

Kong, Y.-Y., Deeks, J. M., Axon, P. R., and Carlyon, R. P. (**2009**). "Limits of temporal pitch in cochlear implants," J. Acoust. Soc. Am. **125**, 1649–1657.

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**, 656–659.

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**, 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**, 2182–2191.

Laback, B., Pok, S.-M., Baumgartner, W.-D., Deutsch, W. A., and Schmid, K. (**2004**). "Sensitivity to interaural level and envelope time differences of two bilateral cochlear implant listeners using clinical sound processors," Ear. Hear. **25**, 488–500.

Laback, B., Zimmermann, I., Majdak, P., Baumgartner, W.-D., and Pok, S.-M. (**2011**). "Effects of envelope shape on interaural envelope delay sensitivity in acoustic and electric hearing," J. Acoust. Soc. Am. **130**, 1515–1529.

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

Laneau, J., Wouters, J., and Moonen, M. (**2006**). "Improved music perception with explicit pitch coding in cochlear implants," Audiol. Neurootol. **11**, 38–52.

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**, 790–802.

McKay, C. M., and Henshall, K. R. (**2010**). "Amplitude modulation and loudness in cochlear implantees," J. Assoc. Res. Otolaryngol. **11**, 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**, 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**, 1777–1785.

McKay, C. M., Remine, M. D., and McDermott, H. J. (**2001**). "Loudness summation for pulsatile electrical stimulation of the cochlea: Effects of rate, electrode separation, level, and mode of stimulation," J. Acoust. Soc. Am. **110**, 1514–1524.

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**, 789–802.

JASA
https://doi.org/10.1121/10.0000610

Milczynski, M., Chang, J. E., Wouters, J., and van Wieringen, A. (**2012**). "Perception of Mandarin Chinese with cochlear implants using enhanced temporal pitch cues," Hear. Res. **285**, 1–12.
Milczynski, M., Wouters, J., and van Wieringen, A. (**2009**). "Improved fundamental frequency coding in cochlear implant signal processing," J. Acoust. Soc. Am. **125**, 2260–2271.
Miller, S. E., Schlauch, R. S., and Watson, P. J. (**2010**). "The effects of fundamental frequency contour manipulations on speech intelligibility in background noise," J. Acoust. Soc. Am. **128**, 435–443.
Monaghan, J. J. M., and Seeber, B. U. (**2016**). "A method to enhance the use of interaural time differences for cochlear implants in reverberant environments," J. Acoust. Soc. Am. **140**, 1116–1129.
Moore, B. C. J., and Carlyon, R. P. (**2005**). "Perception of pitch by people with cochlear hearing loss and by cochlear implant users," in *Pitch*, edited by C. J. Plack, R. R. Fay, A. J. Oxenham, and A. N. Popper (Springer, New York), pp. 234–277.
Morey, R. D. (**2008**). "Confidence intervals from normalized data: A correction to Cousineau (2005)," Tutor. Quant. Methods Psychol. **4**, 61–64.
Oxenham, A. J. (**2012**). "Pitch perception," J. Neurosci. **32**, 13335–13338.
Oxenham, A. J., Micheyl, C., Keebler, M. V., Loper, A., and Santurette, S. (**2011**). "Pitch perception beyond the traditional existence region of pitch," Proc. Natl. Acad. Sci. U.S.A. **108**, 7629–7634.
Pépiot, E. (**2014**). "Male and female speech: A study of mean F0, F0 range, phonation type and speech rate in Parisian French and American English speakers," in *Speech Prosody 7*, edited by N. Campbell, D. Gibbon, and D. Hirst (Trinity College, Dublin, Ireland), pp. 305–309.
Peterson, G. E., and Barney, H. L. (**1952**). "Control methods used in a study of the vowels," J. Acoust. Soc. Am. **24**, 175–184.
Pieper, S. H., and Bahmer, A. (**2019**). "Rate pitch discrimination in cochlear implant users with the use of double pulses and different interpulse intervals," Cochlear Implants Int. **20**, 312–323.
Riss, D., Hamzavi, J.-S., Blineder, M., Honeder, C., Ehrenreich, I., Kaider, A., Baumgartner, W.-D., Gstoettner, W., and Arnoldner, C. (**2014**). "FS4, FS4-p, and FSP: A 4-month crossover study of 3 fine structure sound-coding strategies," Ear. Hear. **35**, e272–e281.
Sayles, M., and Winter, I. M. (**2008**). "Reverberation challenges the temporal representation of the pitch of complex sounds," Neuron **58**, 789–801.
Schouten, J. F. (**1940**). "The residue and the mechanism of hearing," Proc. Koninklijke Nederlandse Akademie Wetenschappen **43**, 991–999.
Shannon, R. V. (**1983**). "Multichannel electrical stimulation of the auditory nerve in man. I. Basic psychophysics," Hear. Res. **11**, 157–189.
Smith, Z. M., Kan, A., Jones, H. G., Buhr-Lawler, M., Godar, S. P., and Litovsky, R. Y. (**2014**). "Hearing better with interaural time differences and bilateral cochlear implants," J. Acoust. Soc. Am. **135**, 2190–2191.
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 Arnoldner, C. (**2020**). "Improving interaural time difference sensitivity using short inter-pulse intervals with amplitude-modulated pulse trains in bilateral cochlear implants," J. Assoc. Res. Otolaryngol.
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. **19**, 301–315.
Stahl, P., Macherey, O., Meunier, S., and Roman, S. (**2016**). "Rate discrimination at low pulse rates in normal-hearing and cochlear implant listeners: Influence of intracochlear stimulation site," J. Acoust. Soc. Am. **139**, 1578–1591.
Stupak, N., Padilla, M., Morse, R. P., and Landsberger, D. M. (**2018**). "Perceptual differences between low-frequency analog and pulsatile stimulation as shown by single- and multidimensional scaling," Trends Hear. **22**, 1–21.
Thakkar, T., Kan, A., Jones, H. G., and Litovsky, R. Y. (**2018**). "Mixed stimulation rates to improve sensitivity of interaural timing differences in bilateral cochlear implant listeners," J. Acoust. Soc. Am. **143**, 1428–1440.
Titze, I. R. (**1989**). "Physiologic and acoustic differences between male and female voices," J. Acoust. Soc. Am. **85**, 1699–1707.
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.
Vandali, A. E., Sucher, C., Tsang, D. J., McKay, C. M., Chew, J. W. D., and McDermott, H. J. (**2005**). "Pitch ranking ability of cochlear implant recipients: A comparison of sound-processing strategies," J. Acoust. Soc. Am. **117**, 3126–3138.
Vandali, A., Sly, D., Cowan, R., and van Hoesel, R. (**2013**). "Pitch and loudness matching of unmodulated and modulated stimuli in cochlear implantees," Hear. Res. **302**, 32–49.
Vandali, A. E., and van Hoesel, R. J. M. (**2011**). "Development of a temporal fundamental frequency coding strategy for cochlear implants," J. Acoust. Soc. Am. **129**, 4023–4036.
Vandali, A. E., and van Hoesel, R. J. M. (**2012**). "Enhancement of temporal cues to pitch in cochlear implants: Effects on pitch ranking," J. Acoust. Soc. Am. **132**, 392–402.
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**, 557–567.
van Hoesel, R. J. M., and Tyler, R. S. (**2003**). "Speech perception, localization, and lateralization with bilateral cochlear implants," J. Acoust. Soc. Am. **113**, 1617–1630.
Vongphoe, M., and Zeng, F.-G. (**2005**). "Speaker recognition with temporal cues in acoustic and electric hearing," J. Acoust. Soc. Am. **118**, 1055–1061.
Williges, B., Jürgens, T., Hu, H., and Dietz, M. (**2018**). "Coherent coding of enhanced interaural cues improves sound localization in noise with bilateral cochlear implants," Trends Hear. **22**, 1–18.
Wilson, B. S. (**2006**). "Speech processing strategies," in *Cochlear Implants: A Practical Guide*, edited by H. Cooper and L. C. Craddock (Whurr, London, UK), 2nd ed., pp. 21–69.
Wilson, B. S., and Dorman, M. F. (**2008**). "Cochlear implants: A remarkable past and a brilliant future," Hear. Res. **242**, 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**, 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**, 67–80.
Xu, L., and Pfingst, B. E. (**2003**). "Relative importance of temporal envelope and fine structure in lexical-tone perception (L)," J. Acoust. Soc. Am. **114**, 3024–3027.
Zierhofer, C. M. (**2003**). "Electrical nerve stimulation based on channel specific sampling sequences," U.S. patent No. 6,594,525.

Perceptual Salience of Temporal Pitch with Consistent and Inconsistent Amplitude Modulation and Short Inter-Pulse Interval Cues

The following manuscript was published in the Psychological and Physiological Acoustics section of the Journal of the Acoustical Society of America[3]:

Lindenbeck, M. J., Majdak, P., Srinivasan, S., & Laback, B. (2023). Pitch discrimination in electric hearing with inconsistent and consistent amplitude-modulation and inter-pulse rate cues. *The Journal of the Acoustical Society of America*, *153*(6), 3268–3280. https://doi.org/10.1121/10.0019452

Author contributions

All authors conceptualized and designed the study. MJL collected the data and performed the statistical analysis. MJL wrote the first draft of the manuscript. All authors contributed to manuscript revision, read and approved the submitted version.

JASA ARTICLE

Pitch discrimination in electric hearing with inconsistent and consistent amplitude-modulation and inter-pulse rate cues[a)]

Martin J. Lindenbeck,[b)] Piotr Majdak, Sridhar Srinivasan, and Bernhard Laback
Acoustics Research Institute, Austrian Academy of Sciences, Wohllebengasse 12-14, A-1040 Vienna, Austria

ABSTRACT:
Users of cochlear implants (CIs) struggle in situations that require selective hearing to focus on a target source while ignoring other sources. One major reason for that is the limited access to timing cues such as temporal pitch or interaural time differences (ITDs). Various approaches to improve timing-cue sensitivity while maintaining speech understanding have been proposed, among them inserting extra pulses with short inter-pulse intervals (SIPIs) into amplitude-modulated (AM) high-rate pulse trains. Indeed, SIPI rates matching the naturally occurring AM rates improve pitch discrimination. For ITD, however, low SIPI rates are required, potentially mismatching the naturally occurring AM rates and thus creating unknown pitch effects. In this study, we investigated the perceptual contribution of AM and SIPI rate to pitch discrimination in five CI listeners and with two AM depths (0.1 and 0.5). Our results show that the SIPI-rate cue generally dominated the percept for both consistent and inconsistent cues. When tested with inconsistent cues, also the AM rate contributed, however, at the large AM depth only. These findings have implications when aiming at jointly improving temporal-pitch and ITD sensitivity in a future mixed-rate stimulation approach. © *2023 Acoustical Society of America*. https://doi.org/10.1121/10.0019452

(Received 28 July 2022; revised 28 April 2023; accepted 1 May 2023; published online 12 June 2023)
[Editor: Li Xu] Pages: 3268–3280

I. INTRODUCTION

The auditory system converts the acoustic signals of multiple sound sources arriving at our ears into mental representations of the individual sources, a process called auditory scene analysis (Bregman, 1990). Successful auditory scene analysis seems to be a prerequisite to selectively attend to the source of interest (e.g., Shinn-Cunningham and Best, 2008). Several cues are exploited in that analysis, among them temporal pitch and interaural time difference (ITD), to which we will collectively refer as timing cues hereafter. For more details see also, for example, Darwin (1997) and Moore and Gockel (2012). Cochlear implants (CIs) are auditory prostheses implanted in cases of severe-to-profound hearing loss or deafness. They were originally designed for conveying basic information required for speech understanding. A landmark development regarding this goal was the continuous interleaved sampling (CIS) stimulation strategy (Wilson *et al.*, 1991) that conveys speech cues by amplitude modulation (AM) of constant-rate pulse train carriers with pulse patterns being interleaved across electrodes to reduce channel interactions (e.g., de Balthasar *et al.*, 2003). In quiet, speech understanding with CIS is robust against replacing the acoustic carrier signal, the temporal fine structure, with pulse trains conveying no fine-structure cues and requires only about four to eight stimulating electrodes to reach asymptotic performance for speech in quiet (Friesen *et al.*, 2001; but see Croghan *et al.*, 2017). CIS focuses on the temporal envelope at the expense of the temporal fine structure of the acoustic signal (Wilson *et al.*, 1991), resulting in rather poor timing-cue sensitivity (pitch: e.g., Gaudrain and Baskent, 2018; ITD: e.g., Laback *et al.*, 2004).

To improve selective hearing with multi-electrode stimulation, three types of stimulation paradigms have been developed. The first type is based on the CIS approach and uses envelope-sharpening for ITD (e.g., Monaghan and Seeber, 2016) and pitch (e.g., Geurts and Wouters, 2001; Laneau *et al.*, 2006; Vandali *et al.*, 2005), respectively. The second type is represented by "partial" high-rate paradigms. They use high pulse rates of at least several hundred pulses per second (pps) on a subset of electrodes to sufficiently sample the speech envelope (Arora *et al.*, 2009; Loizou *et al.*, 2000) and lower rates on the remaining electrodes where the pulse timing is locked to the acoustic fine structure. Examples are peak-derived timing (PDT; van Hoesel and Tyler, 2003), fine structure processing (FSP; Hochmair *et al.*, 2006), FS4 (providing time-locked pulses at low rates on 2–4 apical electrodes; Riss *et al.*, 2014), and the mixed-rate approach (using time-locked pulses on one or a few electrodes somewhere across the array; Churchill *et al.*, 2014; Thakkar *et al.*, 2018). The third type uses time-locked pulses at low rates across the entire array. Examples are fundamental asynchronous stimulus timing (FAST; Hossain and Goldsworthy, 2018; Smith *et al.*, 2014) or peak picking (PP; Williges *et al.*, 2018). Yet, so far, no published approach provided both salient timing cues and sufficient speech information to allow for selective hearing.

[a)]Portions of this work were presented at the 45th Annual Meeting of the German Acoustical Society (DAGA), Rostock, Germany, March 2019, and the 2019 Conference on Implantable Auditory Prostheses, Lake Tahoe, CA.
[b)]Electronic mail: martin.lindenbeck@oeaw.ac.at

 0001-4966/2023/153(6)/3268/13/$30.00

JASA
https://doi.org/10.1121/10.0019452

Most of the above-mentioned clinical high-rate paradigms convey temporal-pitch cues such as the fundamental frequency (F0) of voiced speech primarily by means of encoding of the AM rate in the signal at each electrode. However, sensitivity to such an AM-based temporal pitch is typically poor (Gaudrain and Başkent, 2018), even when presented in a controlled way using a research system (Chatterjee and Oberzut, 2011). Furthermore, AM-based pitch crucially depends on the AM depth (Chatterjee and Oberzut, 2011; Lindenbeck *et al.*, 2020; McKay *et al.*, 1994, 1995; Vandali *et al.*, 2013) and the oversampling factor (McKay *et al.*, 1994, 1995). Furthermore, temporal-pitch sensitivity is limited in pulse rate. On the upper end, it decreases if either the pulse or the AM rate exceed approximately 200 pps (Kong *et al.*, 2009; Lindenbeck *et al.*, 2020, Townshend *et al.*, 1987; Vandali *et al.*, 2013). This upper rate limitation is similar to a rate limitation observed for the sensitivity to ITDs (Ihlefeld *et al.*, 2015; Laback *et al.*, 2015; Laback *et al.*, 2007; Majdak *et al.*, 2006). On the lower end, pitch perception is limited to frequencies above 30 Hz in normal hearing (Krumbholz *et al.*, 2000; Pressnitzer *et al.*, 2001). It is reasonable to assume such a lower rate limitation also for electric hearing (Chatterjee and Oberzut, 2011, Kong *et al.*, 2009, and Lindenbeck *et al.*, 2020), with frequencies below 60 Hz producing a buzz rather than a pitch (McKay and Carlyon, 1999). Taken together, there seems to be a "sweet spot" for temporal pitch with CIs for rates ranging between 60 and 200 pps.

There has been a debate on whether more than one temporal-pitch cue can be conveyed via the same cochlear location. In electric hearing, a model for pitch sensitivity to high-rate AM pulse trains (alternating between two currents with integer carrier-to-AM-rate ratios) was proposed (McKay *et al.*, 1994, 1995). It assumed that the pitch percept of such stimuli would be matched to that of an unmodulated pulse train with a rate equal to a weighted average of carrier and AM rate with the AM depth being the weighting factor. The model reasonably well explained the psychophysical pitch-matching data. Furthermore, the model was based on the current-dependent recruitment of auditory-nerve neurons but was unspecific whether the weighting of carrier and AM rate was sensory (listeners only perceive the compound pitch cue) or cognitive (listeners perceive both carrier and AM-pitch cue). In a subsequent study by Carlyon (1997) utilizing a simulation of electric hearing, normal-hearing listeners were not able to discriminate between two concurrent band-limited and inharmonically related acoustic pulse trains based on differences in F0 alone. Yet, in electric hearing with AM pulse trains providing both a salient carrier-rate cue and a salient harmonically related AM-rate cue, McKay and Carlyon (1999) showed that CI listeners were able to extract both rate cues from a signal. Furthermore, with simulated electric hearing, they showed that the perceptual weighting of these cues varied systematically with the AM depth. These results are further supported by evidence from normal-hearing studies suggesting that the pitch of three concurrent and harmonically related complex tones can be extracted simultaneously based on temporal cues alone (Graves and Oxenham, 2019). Taken together, these studies suggest that at least two to three concurrent purely temporal pitch cues are available to the auditory system.

Recently, Lindenbeck *et al.* (2020) showed that inserting pulses with short inter-pulse intervals (SIPI) after carrier pulses at AM peaks (the so-called SIPI pulses) improves single-electrode temporal-pitch sensitivity, particularly at low AM depths (< 0.5). SIPI pulses also improve ITD sensitivity both with unmodulated and AM high-rate pulse trains (Srinivasan *et al.*, 2020; Srinivasan *et al.*, 2018) with the highest ITD sensitivity in case of unmodulated pulse trains being at SIPI rates below 100 pps (Srinivasan *et al.*, 2018). Surprisingly, for AM pulse trains and a fixed SIPI rate of 63 pps, sensitivity did not depend on the F0, suggesting that ITD sensitivity benefits from low SIPI rates regardless of the F0 (Srinivasan *et al.*, 2020). This leads to a pronounced difference in the optimal configuration required to achieve improvements in sensitivity in both temporal pitch and ITD (Lindenbeck *et al.*, 2020; Srinivasan *et al.*, 2020; Srinivasan *et al.*, 2018). For pitch and AM rates in the range of naturally occurring speech F0s (e.g., Peterson and Barney, 1952), matching SIPI and AM rate ("full-rate" SIPI condition) yielded larger and more homogeneous improvements in sensitivity across listeners compared to the condition without SIPI pulses. Mismatched AM and SIPI rates, that is, a SIPI rate of half the AM rate ("half-rate" SIPI condition), yielded lower performance compared to the condition without SIPI pulses. Despite that, it is still unclear whether the SIPI approach is able to improve timing-cue sensitivity in a configuration of mismatched AM and SIPI rates.

To this end, in our study, we investigated the ability to discriminate pitch with single-electrode AM high-rate pulse trains supplemented with SIPI pulses. Hence our stimuli contained two rate cues that are potentially important for pitch perception, namely, AM and SIPI-based cues. We speculated the CI listeners rely more on the SIPI-rate cue than on the AM rate cue based on the observation that half-rate SIPI discrimination performance was more similar to the performance obtained with unmodulated pulse trains with rates matching the SIPI rate, rather than with rates matching the AM rate (cf. Lindenbeck *et al.*, 2020).

To find out whether SIPI rate or AM rate determines the pitch of high-rate AM pulse trains with SIPI pulses, we measured single-electrode pitch-discrimination sensitivity with several combinations of *mismatched* and matched AM and SIPI rates, covering a range of AM rates and a range of SIPI rates. Similar to McKay and Carlyon (1999), we assumed that the AM- and the SIPI-rate cues are extracted separately and weighted by the auditory system of the CI listener. In a two-interval discrimination task, we provided two concurrent rate trajectories of SIPI and AM in each trial, quantified by the task-difficulty parameters ΔSIPI and ΔAM, respectively. Our first hypothesis was that the SIPI pulses provide the dominant rate cue. To assess this hypothesis, we analyzed the effect sizes of AM- and SIPI-related stimulus parameters. We predicted a perceptual dominance of the

https://doi.org/10.1121/10.0019452 JASA

SIPI-based pitch cue that is reflected in larger effect sizes for SIPI-related parameters. Furthermore, we hypothesized that the AM cue becomes more salient with increasing AM depth (e.g., Geurts and Wouters, 2001; Lindenbeck *et al.*, 2020) and tested two depths of 0.1 and 0.5. We hypothesized that the relative contribution of AM cues is higher at the higher depth and that this is reflected in differences between the effect sizes across AM depths.

II. METHODS

A. Listeners

Five post-lingually deafened CI listeners participated. Their etiological details are listed in Table I. All participants had 12-channel CIs manufactured by MED-EL Corp. (Innsbruck, Austria). CI17, CI24, and CI117 are bilaterally implanted and were tested with their preferred ear, which always corresponded to the ear implanted first. CI17, CI18, and CI24 were tested shortly after participating in Lindenbeck *et al.* (2020). CI116 and CI117 were tested during the COVID-19 pandemic. Safety measures were put in place to provide the highest possible safety at the given stage of the pandemic. All participants were paid an hourly wage for their participation. There were no additional payments for pandemic risks. All participants gave informed consent before conducting the study. CI116 and CI117 were informed about the safety measures when first contacted for participation.

B. Stimuli and apparatus

All stimuli were single-electrode 2000-pps pulse trains with an AM and with a single SIPI pulse inserted at every AM peak. The AM shape was a full-wave rectified sinusoid and the rate *after* rectification was defined as the AM rate. The stimuli were presented to electrode eight, for which the carrier rate approximately matched the place (cf. Baumann and Nobbe, 2004), and had a nominal duration of 600 ms including linear 150-ms onset and offset ramps (the ramps did not contain SIPI pulses). Relative to the AM, SIPI pulses were inserted at the peak of either every AM period (full-rate SIPI condition), every other AM period (half-rate SIPI condition), or every fourth AM period (quarter-rate SIPI condition). Figure 1(A) shows a few cycles of an example stimulus in which the SIPI rate is half the AM rate (i.e., a half-rate-SIPI stimulus). It was always ensured that each AM period was sampled by the pulse train in the same way, i.e., only a limited range of AM rates that are integer submultiples of the carrier rate were permitted. Thereby, beat frequency distortion was avoided (Goldsworthy *et al.*, 2022) and sufficient oversampling (McKay *et al.*, 1994) was ensured [the highest AM/SIPI rate was 250 Hz/pps, cf. Fig. 1(B)]. The time gap between carrier and SIPI pulses was 60 μs, which amounts to 12% of the carrier inter-pulse interval of 500 μs (cf. Lindenbeck *et al.*, 2020). The carrier pulse trains consisted of biphasic pulses with a phase duration of 26.7 μs for C40+ implants and 27.5 μs for Pulsar and Synchrony implants. The implants were used in monopolar mode. The stimuli were otherwise identical to the SIPI stimuli used in Lindenbeck *et al.* (2020). The particular choice of stimulus conditions is motivated in Sec. II E.

The stimulus amplitudes were defined in linear current units. To estimate the dynamic range for each CI listener, threshold and maximum comfortable level were determined by the experimenter with a fitting procedure employing 2000-pps pulse trains without an AM.

The implants were connected to a personal computer via the Research Interface Box II (developed at the Institute of Ion Physics and Applied Physics, Leopold-Franzens-University, Innsbruck, Austria) which permits precise definition of pulse parameters and bypassing of the clinical sound processors. The stimuli were generated in MATLAB (The MathWorks Inc., Natick, MA, USA) and the experiment was controlled using the custom-written software EXPSUITE (Sourceforge, 2023).

C. Loudness balancing

With stimuli as in Fig. 1(A), changes in SIPI rate affect loudness (Lindenbeck *et al.*, 2020), comparable to the pulse-rate effect for unmodulated pulse trains (McKay *et al.*, 2001). Hence, we loudness-balanced all seven experimental stimuli as shown in Fig. 1(B). Four of these stimuli (i.e., the 125-Hz half-rate, 125-Hz full-rate, 250-Hz half-rate, and 250-Hz full-rate stimulus) were formally balanced using the same adaptive 3-up 1-down/1-up 3-down staircase procedure with a two-interval, two-alternative forced-choice task as described by Lindenbeck *et al.* (2020), asking listeners to indicate whether the second target sound was louder or softer than the preceding reference. From these balanced

TABLE I. Details for the participating CI listeners. If the onset of deafness could not be attributed to a specific date (e.g., in case of progressive hearing loss), "age at onset of deafness" may also refer to the onset of profound hearing loss.

				Tested ear			
Listener	Sex	Age (yr)	Etiology	Side	Age at onset of deafness (yr)	CI experience (yr)	MED-EL implant type
CI17	female	71	Idiopathic	R	41	13	Synchrony
CI18	male	60	Progressive	R	47	13	Pulsar
CI24	female	55	Progressive	L	41	13	C40+
CI116	male	57	Progressive	R	55	2	Synchrony
CI117	male	40	Progressive	R	19	18	C40+
M (*SD*)	—	56.6 (11.1)	—	—	40.6 (13.4)	11.8 (5.9)	—

JASA
https://doi.org/10.1121/10.0019452

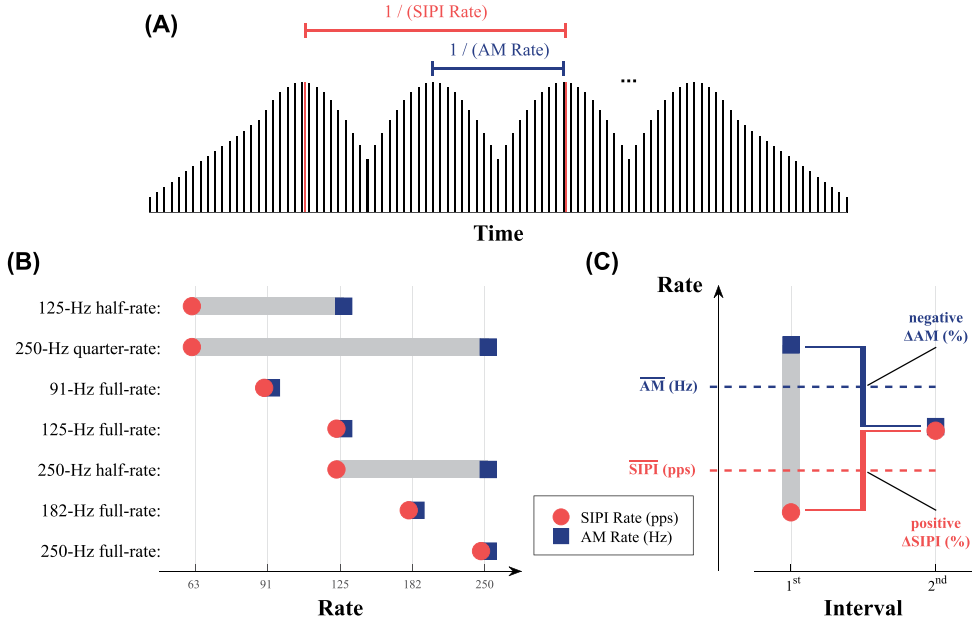


FIG. 1. (Color online) Stimulus specification: (A) Stimulus example in which the SIPI rate is half the AM rate. Negative pulse phases are omitted and only a few AM cycles of the actual stimulus are shown (indicated by the dots). (B) Combinations of AM and SIPI rate, where stimuli are sorted in ascending order of SIPI rate from top to bottom and denoted by the AM rate (in Hz) and the SIPI rate relative to the AM rate [(full-rate [SIPI rate = AM rate], half-rate [SIPI rate = AM rate/2], and quarter-rate [SIPI rate = AM rate/4], cf. Lindenbeck *et al.*, 2020. (C) An example of an inconsistent-cue condition (i.e., 250-Hz half-rate) in the two-interval discrimination task. Task-sensitivity parameters: the AM rate decreases from the first to the second interval ($\Delta AM < 0$), the SIPI rate increases ($\Delta SIPI > 0$). The averaged (absolute) AM rate ($\overline{AM}$) is higher than the averaged SIPI rate ($\overline{SIPI}$).

amplitudes, the amplitudes for the three other stimuli (i.e., the 250-Hz quarter-rate, the 91-Hz full-rate, and the 182-Hz full-rate stimulus) were extrapolated and double-checked by informally asking the subjects to indicate balanced loudness on a visual "same-different" loudness scale. In addition, 3% dynamic-range level roving was applied independently across intervals for all following pitch tests.

D. Pretest

To assess whether listeners were sensitive to the SIPI and AM rates investigated in our study, a pitch-sensitivity pretest preceded the main pitch-discrimination experiment. A two-interval two-alternative forced-choice task was combined with the method of constant stimuli employing a fixed set of frequency differences (FDs, expressed in % *re* the lower of the two frequencies) for 2000-pps AM stimuli either without SIPI pulses (AM condition) or with SIPI pulses inserted at the AM rate (so-called AM-plus-SIPI condition). The pretest was conducted for two AM rates of 125 and 250 Hz. Note that these rates corresponded to the geometric average of the rates to be discriminated at a given FD (Kreft *et al.*, 2010). Assuming that temporal-pitch sensitivity improves with increasing AM depth, the AM depth of 0.1 was tested only. Three FDs were tested, and, because of the stimulus properties, these three FDs differed between AM rates: For 125 Hz, FDs of 6%, 20%, and 46% were used. For 250 Hz, FDs of 13%, 29%, and 57% were used. Listeners CI17, CI18, and CI24 were tested as part of Lindenbeck *et al.* (2020), who used 100 repetitions per condition. Listeners CI116 and CI117 were tested in our study and with 50 repetitions per condition. Apart from this difference, our procedure was the same as in Lindenbeck *et al.* (2020).

All listeners received written instructions and indicated whether the pitch of the second stimulus was higher or lower than the pitch of the first stimulus. They were furthermore instructed to always focus on the dominant pitch. Within each trial, stimuli were separated by a gap of 490 ms and a new trial was played 310 ms after the response. Trial-by-trial feedback was provided. Per condition, both orders of stimuli were presented equally often and the percent-correct results for both temporal orders of stimuli were converted to d' scores (Klein, 2001) to account for bias towards one order of stimuli. With 50 repetitions, d' scores in the range of ± 0.52 indicate performance that is not significantly different from chance (95% confidence interval, Monte Carlo simulation). To prevent the d' score from reaching infinity, 100% scores were set to 99%, limiting the d' score to 3.3. To familiarize the listeners with the task, a brief training session preceded the pretest. Breaks were allowed at any time.

After data collection, we estimated temporal-pitch sensitivity for a criterion indicating whether listeners performed better than chance ($d' = 0.52$, 50 repetitions).[1] The estimates were calculated by linearly interpolating the two d' scores surrounding the criterion on the psychometric function of FD.

Figure 2 shows the results of the sensitivity estimation. Note that the lowest possible estimates are 6% for the rate

https://doi.org/10.1121/10.0019452 JASA

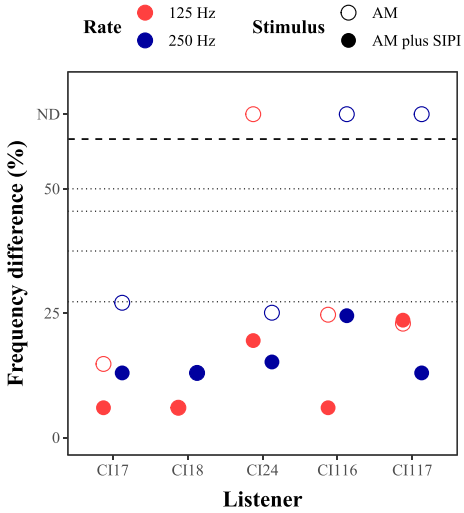


FIG. 2. (Color online) Pitch-discrimination sensitivity in the pretest. Frequency differences (FD, in %) are shown for the AM depth of 0.1, two rates (red: 125 Hz; blue: 250 Hz), and two AM conditions (open symbols: "AM"; filled symbols: "AM plus SIPI"). The criterion was a d' of 0.52 indicating performance beyond chance for $N = 50$ repetitions. Symbols above the dashed line show non-discriminable conditions. Dotted lines show frequency differences used in the main pitch-discrimination experiment. For CI18, AM, and AM-plus-SIPI sensitivity were identical.

of 125 Hz and 13% for the rate of 250 Hz, respectively. The CI listeners were—except for three conditions for which sensitivity was not detectable (ND)—sensitive enough to *not* have to guess even for the smallest FD of the main experiment (i.e., all data points fall below the dotted line at the frequency difference of 27%). In general, listeners were more sensitive when tested for the rate of 125 Hz and in the AM-plus-SIPI conditions, which is in line with the results from Lindenbeck *et al.* (2020).

E. Main experiment

While in the pretest SIPI and AM rate were consistently changed within a trial, the main experiment was concerned with inconsistent within-trial changes of SIPI and AM rate, respectively. The same task and instructions as in the pretest were used, but no feedback was provided.[2] Again, the method of constant stimuli was used to compare all 21 possible pairs of the seven stimuli [cf. Fig. 1(B)] in both temporal orders of each pair. Each pair was tested 50 times per AM depth, resulting in a total of 2100 trials (presented in completely randomized order). Listeners were allowed to take breaks at any time. Responses with d' scores greater than zero indicate perception of ascending pitch across intervals, whereas d' scores below zero indicate perception of descending pitch.

1. Conditions

Figure 1(B) provides an overview of the seven stimuli that were compared pairwise to each other. The stimulus rate (SIPI with red circles and AM with blue squares) is indicated along the horizontal axis. The stimuli are labeled by their AM rate (91, 125, 182, and 250 Hz) and the SIPI rate relative to the AM rate (full-rate, half-rate, and quarter-rate). In Fig. 1(B), they are sorted by ascending SIPI rate.

The stimuli were selected to study the relative contributions of SIPI and AM-rate cues across practically relevant ranges of rates for those two cues, based on preceding studies (Lindenbeck *et al.*, 2020; Srinivasan *et al.*, 2020, 2018). Four of the stimuli (125 and 250-Hz AM rates combined with full-rate and half-rate SIPI pulses) were used by Lindenbeck *et al.* (2020). 91- and 182-Hz stimuli have intermediate rates close to the geometric averages of 63 and 125 Hz or 125 and 250 Hz, respectively.[3]

The pairwise combination of stimuli yielded 21 conditions per AM depth which could be clustered in three groups based on their rate-cue *changes*. First, in a "*consistent*" condition (e.g., a 125-Hz full-rate-SIPI stimulus followed by a 250-Hz full-rate-SIPI stimulus or vice versa; 10 conditions in total), AM and SIPI rates both change in the same direction, creating coherent SIPI and AM-rate trajectories. Thus, these conditions are broadly similar to Lindenbeck *et al.* (2020). Second, in an "opposed" condition (e.g., a 125-Hz half-rate-SIPI stimulus followed by a 91-Hz full-rate-SIPI stimulus or vice versa; 5 conditions in total), the SIPI rate changed in one direction while the AM rate changed in the other direction. Third, in a "single" condition (e.g., a 125-Hz full-rate-SIPI stimulus followed by a 250-Hz half-rate-SIPI stimulus or vice versa; 6 conditions in total), either SIPI or AM rate remained constant across intervals, creating one changing and one constant rate trajectory. However, because in both "opposed" and "single" conditions the rate trajectories were incoherent, these two conditions will collectively be referred to as *inconsistent* hereafter.

Figure 1(C) shows an example for an "opposed" condition: the SIPI rate increases (ascending trajectory) while the AM rate decreases (descending trajectory). It further shows the four parameters that were defined to *jointly* specify each cue-change condition. The rate difference is represented by the absolute values of the SIPI-rate difference (ΔSIPI) and the AM-rate difference (ΔAM), respectively, which are defined as Weber fractions (in %). The direction of the rate trajectory is represented by the sign of ΔSIPI and ΔAM: positive for ascending and negative for descending rate. Table II further summarizes the rate differences and

TABLE II. Summary of cue-change conditions for one of the two possible orders of stimuli: Rate difference (i.e., ΔSIPI and ΔAM, respectively) and corresponding rate trajectory from stimulus one to stimulus two (ascending [↗], constant [→], or descending [↘]). To obtain the summary for the second possible order of stimuli, mutually exchange < with > and ↗ with ↘.

	Rate difference		Rate trajectory	
Cue-change condition	SIPI	AM	SIPI	AM
Consistent	>0	>0	↗	↗
Opposed	>0	<0	↗	↘
Single	$>0\,/\,0$	$0\,/\,>0$	↗ / →	→ / ↗

JASA
https://doi.org/10.1121/10.0019452

trajectories for all three types of conditions. Note that both directions of rate change (low to high and vice versa) were tested with both possible rate-difference signs (i.e., $\Delta > 0$ and $\Delta < 0$, respectively). For the sake of simplicity, the descriptions hereafter refer only to one (arbitrarily chosen) order of intervals.

In Fig. 1(C), the absolute rate is represented as the geometric average of the rates across both intervals, with $\overline{\text{SIPI}}$ (in pps) and $\overline{\text{AM}}$ (in Hz) for SIPI and AM rates, respectively. The example in the figure shows an ascending SIPI trajectory ($\Delta\text{SIPI} > 0$) combined with a descending AM trajectory ($\Delta\text{AM} < 0$). In this example trial, $\overline{\text{AM}}$ is larger than $\overline{\text{SIPI}}$. Due to the nature of the SIPI approach, extra pulses were inserted with a rate that was never larger than the AM rate (i.e., $\overline{\text{AM}} \geq \overline{\text{SIPI}}$ for all conditions).

2. Statistics

For each AM depth, each condition is *simultaneously* influenced by the two rate-difference parameters ΔSIPI and ΔAM and the two rate-sensitivity parameters $\overline{\text{SIPI}}$ and $\overline{\text{AM}}$. Consequently, we employed linear-regression analyses with a significance level of 5% to disentangle the effects of these parameters. In the models, we included ΔSIPI, ΔAM, $\overline{\text{SIPI}}$, and $\overline{\text{AM}}$ as continuous predictors.[4] Due to the within-subjects design of the experiment, we always included the Listener as a categorical predictor to partition out the between-subjects variance (Bland and Altman, 1994, 1995), but never report its effect since it is not subject to interpretation. To this end, we hereafter refer to the models as repeated-measures regression analyses (rmREGs).

The goodness of fit of the rmREGs was quantified with the adjusted R^2 (R_a^2). To provide a measure of uncertainty for the fits, we provide bootstrapped[5] 95% R_a^2 confidence intervals (R package "boot"; Canty and Ripley, 2022; cf. Ohtani, 2000). Separate rmREGs were conducted per AM depth to assess its effect by comparing predictors across rmREGs. All statistical analyses were performed using R 4.2.0 (R Core Team, 2022).

Our hypothesis evaluations were based on relative differences between AM- and SIPI-related standardized regression coefficients β both within and across rmREGs. β quantifies the relation between the condition parameters and temporal-pitch sensitivity independently of the parameter ranges. Thus, the strength of the relation can be classified as either small ($\beta \geq 0.1$), medium ($\beta \geq 0.3$), or large ($\beta \geq 0.5$, e.g., Cohen, 1992). To provide estimates of effect-size uncertainty, we provide bootstrapped 95% β confidence

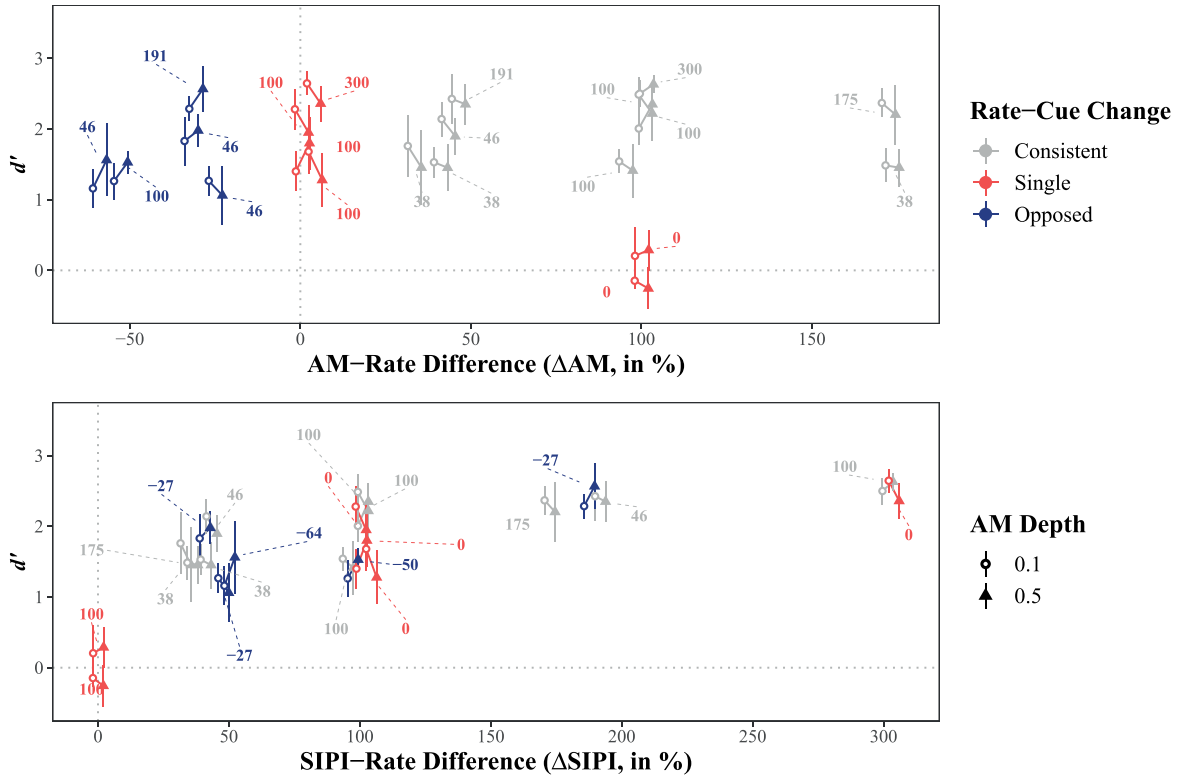


FIG. 3. (Color online) Relation between pitch-discrimination *d'* scores and rate differences. Each condition is a unique combination of AM-rate difference (ΔAM, –64% to 175%), SIPI-rate difference (ΔSIPI, 0% to 300%), average AM rate ($\overline{\text{AM}}$, 107 to 250 Hz), average SIPI rate ($\overline{\text{SIPI}}$, 63 to 213 pps), and AM depth (0.1 and 0.5). Both panels are based on the *same* data. The symbol shapes denote the two AM depths and the symbol colors denote the types of cue-change conditions (cf. Sec. II E 1). Lines connect the same conditions but for the two tested AM depths. Numbers indicate the rate of the cue that is *not* shown on the abscissa. The data are averaged across CI listeners and error bars show normalized standard errors (Cousineau, 2005; Morey, 2008). *Top panel*: ΔAM; annotation: ΔSIPI. *Bottom panel*: ΔSIPI; annotation: ΔAM.

https://doi.org/10.1121/10.0019452 JASA

intervals. Note that the confidence intervals link effect size and statistical significance in that the intervals for significant effects do not include a β of 0.

III. RESULTS

A. Visual inspection

Figure 3 shows the dependency of pitch-discrimination d' scores on the two task-difficulty parameters, that is, the rate differences ΔAM (top row) and ΔSIPI (bottom row). Both panels show the *same* data but with a different rate difference on the abscissa. The annotations show the rate differences *not* shown in the abscissa. The lines link conditions differing only in AM depth. These two panels provide an easy-to-access overview for the relative contributions of ΔAM and ΔSIPI: If discrimination were solely based on ΔAM, a d' of 0 would be observed at a ΔAM of 0 (and vice versa for ΔSIPI).

For ΔAM (top panel), the three types of cue-change conditions created clearly distinguishable clusters (each cluster is shown in a separate color). In the "consistent" condition (gray), performance was better than guessing ($d' > 0.5$) and did not seem to systematically depend on the size of ΔAM. In the "single" condition (red), the performance clustered around two values of ΔAM: At a ΔAM of 0% (same AM rates, but different SIPI rates), d' scores span the entire supra-threshold ($d' > 1$) range. At a ΔAM of 100% (one octave AM-rate difference, but the same SIPI rates), d' scores centered around the threshold of guessing ($d' = 0$). Both observations combined suggest that the performance did not systematically depend on the size of ΔAM. In the "opposed" condition (blue), all d' scores were *positive*, despite the ΔAM being *negative* (i.e., < 0). This suggests that ΔAM did not have a large effect on the d' scores because if CI listeners were relying on the AM-based pitch cue, d' scores would have also been negative. The impression is further supported by the d' scores increasing with ΔSIPI for each ΔAM.

For ΔSIPI (bottom panel), we were not able to observe any clear clustering by type of cue-change condition. In the "consistent" condition (gray), performance was better than guessing, with the pattern similar to that observed in the top panel, but with a stronger dependency on the size of ΔSIPI. In the "single" condition (red), a ΔSIPI of 0% (equal SIPI rates, but ΔAM > 0) yielded d' scores at the threshold of guessing. With increasing ΔSIPI (but equal AM rates), the d' scores also increased. In the "opposed" condition (blue), d' scores were positive and increased with ΔSIPI, despite the sometimes negative ΔAM. The impression that ΔSIPI did have a considerable effect on the d' scores is further supported by the ΔAMs annotated in the figure: For each ΔSIPI, the d' scores did not vary systematically with ΔAM.

Two further observations can be made. First, the effect of the AM depth (observed across both panels) was small, with changes in d' scores being well within the standard errors. Second, the combination of ΔAM of 100% and ΔSIPI of 0% yielded d' scores roughly at chance rate, clearly representing an "outlier" condition (for both AM depths) that required extra treatment in the following statistical analysis.

In summary, our visual inspection raised evidence for ΔSIPI dominating the pitch discrimination, with little perceptual contribution of ΔAM. However, note that the rate-sensitivity parameters $\overline{\text{AM}}$ and $\overline{\text{SIPI}}$ were not inspected so far.

B. Statistical analysis

The statistical analysis was split between consistent and inconsistent conditions. The rationale was that inconsistent conditions were more suitable to assess the weighting of AM-rate and SIPI-rate cues while consistent conditions were more suitable to assess rate sensitivity *per se* and to link the present data to the previous study by Lindenbeck *et al.* (2020). The relations which are most informative regarding our hypothesis are included in the figures below.

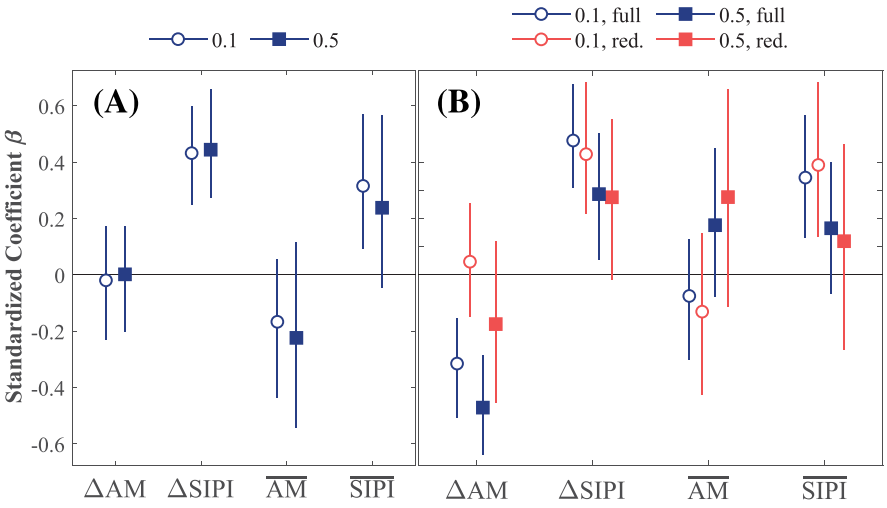


FIG. 4. (Color online) Standardized regression coefficient β alongside 95% confidence intervals separately per AM depth (markers) and for AM-rate difference (ΔAM), SIPI-rate difference (ΔSIPI), average AM rate ($\overline{\text{AM}}$), and average SIPI rate ($\overline{\text{SIPI}}$), for (A) consistent conditions and (B) inconsistent conditions. For the latter, the reduced data sets (in red) excluded two highly influential "single" conditions (referred to as "outlier" condition). Note that parameter ranges differed slightly between consistent and inconsistent conditions.

JASA
https://doi.org/10.1121/10.0019452

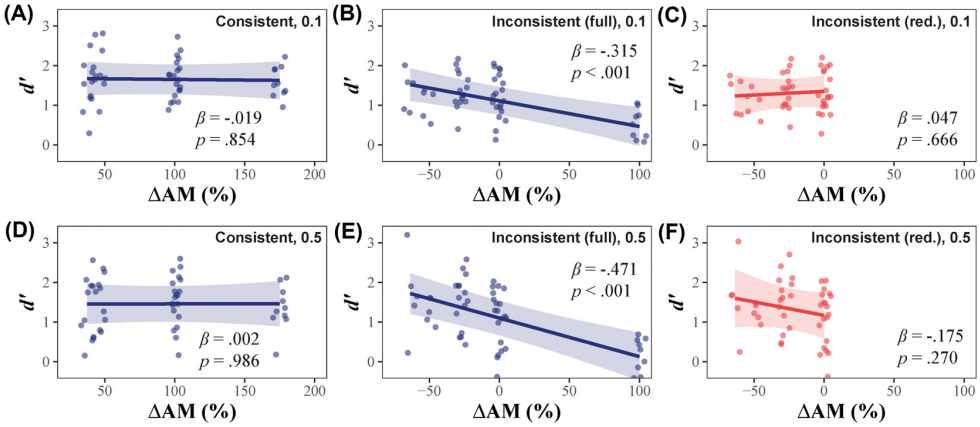


FIG. 5. (Color online) Relation (i.e., partial correlation) between AM-rate difference (ΔAM) and temporal-pitch sensitivity. Per panel, data set [consistent and inconsistent (both "single" and "opposed")] and AM depth (0.1 and 0.5), standardized regression coefficient β, and statistical significance are denoted. Panels (A), (B), (D), and (E) (all in blue) show results for the full data set, and panels (C) and (F) (both in red) show data for the reduced data set.

For completeness, the other relations are provided in the supplementary material.[6]

1. Consistent conditions

Two rmREGs were conducted on the d' scores (one per AM depth). Each rmREG contained as predictors the condition parameters ΔAM (38% to 175%), ΔSIPI (38% to 300%), $\overline{\text{AM}}$ (107 to 213 Hz), and $\overline{\text{SIPI}}$ (88 to 213 pps). Per rmREG, all $N = 50$ data points available (10 per CI listener) were included. The overall rmREG goodness-of-fit statistics are summarized in Table SI of the supplementary material, the resulting parameters are summarized in Table SII, and the β coefficients representing the effect size of each cue are shown in Fig. 4(A).

The differences between AM depths were negligible. Across both depths, ΔAM did not affect the d' scores [cf. Figs. 5(A) and 5(D)]. In contrast, the d' scores improved with ΔSIPI [cf. Fig. 6(A) and 6(D)]. Furthermore, the d' scores tended to worsen with increasing $\overline{\text{AM}}$ [cf. Figs. S1(A) and S1(D)] but improved with increasing $\overline{\text{SIPI}}$ [cf. Fig. S2(A) and S2(D)]. However, note that there was considerable variability in the effects of the rate-sensitivity parameters as expressed by the β confidence intervals, particularly for the large AM depth.

2. Inconsistent conditions

Two rmREGs were conducted on the d' scores (one per AM depth). Each rmREG contained as predictors the condition parameters ΔAM (–64% to 100%), ΔSIPI (0% to 300%), $\overline{\text{AM}}$ (107 to 250 Hz), and $\overline{\text{SIPI}}$ (63 to 177 pps). Per rmREG, all $N = 55$ data points available (11 per CI listener) were included. The overall rmREG goodness-of-fit statistics are summarized in Table SI, the parameters are summarized in Table SIII, and the β coefficients are shown in Fig. 4(B) in blue ("full").

The SI d' scores *worsened* with increasing ΔAM, however, this effect tended to be larger for the larger AM depth [cf. Fig. 5(B) and 5(E)]. Furthermore, the scores improved with increasing ΔSIPI, however, that effect tended to be smaller for the larger AM depth [cf. Figs. 6(B) and 6(E)].

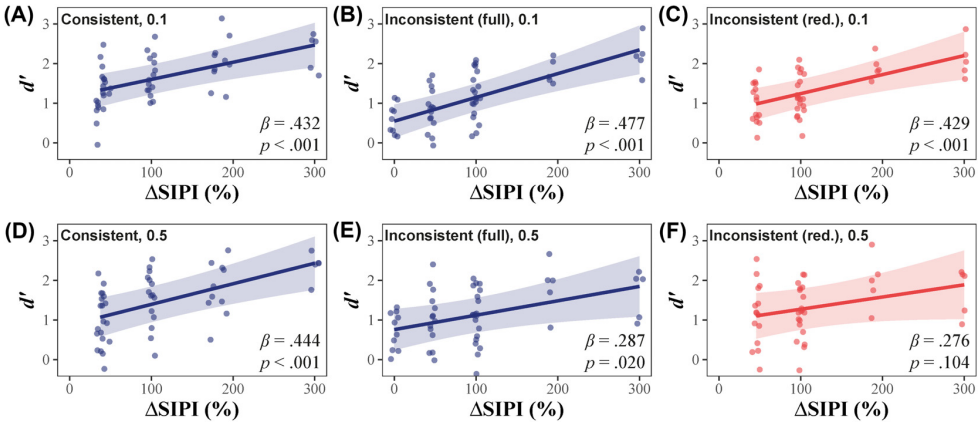


FIG. 6. (Color online) Relation between SIPI-rate difference (ΔSIPI) and temporal-pitch sensitivity. All other aspects as in Fig. 5.

https://doi.org/10.1121/10.0019452 JASA

With respect to average rates, different from consistent conditions, the d' scores tended to worsen slightly with increasing $\overline{\text{AM}}$ for the AM depth of 0.1 but tended to improve with increasing $\overline{\text{AM}}$ for the AM depth of 0.5. Furthermore, the d' scores improved with increasing $\overline{\text{SIPI}}$ for both AM depths, however, that effect was smaller for the larger AM depth.

The negative relation between ΔAM and the d' scores is counterintuitive and was not predicted. To better understand the origin of this effect we revisited the experimental setup for the inconsistent conditions. Indeed, two "single" conditions (in red in Fig. 3) with a ΔAM of 100%—combined with a ΔSIPI of 0%—could be considered outliers in the setup that might have been the main driver of the negative effect. Note that with "outlier," we mean that the data are valid but, due to the inherently unbalanced experimental design,[7] might have had a high leverage on the rmREG results.

To investigate the contribution of these outlier conditions to the negative relation between ΔAM and the d' scores, we repeated the rmREG analysis of the inconsistent conditions with a reduced data set that excluded the two outlier conditions. Per rmREG, all $N = 45$ data points available (9 per CI listener) were included. The goodness-of-fit results are provided in Table SI. The parameter results are summarized in Table SIV, and the associated β coefficients are visualized in Fig. 4(B) in red.

Compared to the full data set, with the reduced data set the negative relation between ΔAM and the d' vanished for the AM depth of 0.1 [cf. Fig. 5(C)], and became considerably smaller (and non-significant) for the AM depth of 0.5 [cf. Fig. 5(F)]. The effects ΔSIPI, $\overline{\text{AM}}$, and $\overline{\text{SIPI}}$ changed only negligibly as compared to the full data set [cf. Figs. 6(C) and 6(F), S1(C) and S1(F), and S2(C) and S2(F)].

Taken together, with the reduced data set, the effect of ΔAM was similar between consistent and inconsistent conditions for the AM depth of 0.1 and lay in between that for the consistent conditions and that for the full inconsistent data set for the AM depth of 0.1. This suggests that, when excluding the outliers, the SIPI rate was the only relevant rate cue for the AM depth of 0.1 even in inconsistent conditions. For the AM depth of 0.5, there was a trend towards a larger contribution of the AM-rate cue. Combining the results from the reduced data set with those from the full data set further suggests that the negative effect of ΔAM in the full data set was actually a combination of two underlying perceptual effects. In particular, the primary effect seemed to have been caused by the dominance of the SIPI-rate cue, causing a stark reduction of sensitivity if the SIPI-rate difference is zero. However, a residual effect seemed to remain even when removing this condition from the analysis, especially for the larger AM depth.

Furthermore, in Fig. 4(B), the β confidence intervals were enlarged for the reduced as compared to the full data set. This was the case particularly for the AM depth of 0.5. While this increase might in principle be due to the smaller reduced data set *per se*, increased confidence intervals for the AM depth of 0.5 compared to 0.1 were also observed for the rate-sensitivity parameters with the (full) consistent data set (for which the sample size was matched for both AM depths, cf. Sec. III B 1). Hence, the increased confidence intervals for the reduced data and the AM depth of 0.5 might also result from increased across-listener variability.

IV. DISCUSSION

To investigate the hypothesis that the SIPI rate dominates temporal-pitch sensitivity when AM-rate and SIPI-rate cue are inconsistent, we compared the effects of rate differences and average rates observed for inconsistent (i.e., "single" and "opposed") conditions with those observed for consistent conditions. Results were also compared across two AM depths to test the hypothesis that the AM cue becomes more salient with increasing AM depth.

A. Effect of rate difference

We quantified the rate differences by ΔSIPI and ΔAM. In theory, increasing Δ should result in increasing d' scores. For a given AM depth, we hypothesized that differences between the effect sizes (i.e., the standardized regression coefficients β from the rmREGs) of ΔAM and ΔSIPI reflect the relative weightings of AM-rate and SIPI-rate cue in the rate-discrimination judgments. At low AM depths, SIPI pulses improve temporal-pitch sensitivity across various F0s (Lindenbeck *et al.*, 2020), thus, in the present study, we expected both an increasing performance with increasing Δ (i.e., a positive effect of Δ) and a larger effect size (i.e., larger $|\beta|$) for ΔSIPI than for ΔAM.

For consistent conditions and at both AM depths, we found no effect for ΔAM but a medium (highly significant) positive effect of ΔSIPI. Note that a medium effect is "likely to be visible to the naked eye of a careful observer" (see Cohen, 1992, p. 156). This suggests that the SIPI-rate cue dominated rate-discrimination sensitivity (regardless of the AM depth). Our rate differences were largely supra-threshold, that is, it should have been rather easy to discriminate them (cf. the pretest results shown in Fig. 2). In contrast, in Lindenbeck *et al.* (2020), AM-only and AM-plus-SIPI sensitivity was similar when tested at the AM depth of 0.5 and with a rate difference near the rate-discrimination threshold. It is thus reasonable to assume that both AM-rate and SIPI-rate cues were roughly similarly salient at this depth in Lindenbeck *et al.* (2020). Interestingly, in the present study, we did not find an effect of ΔAM, not even for this depth, suggesting that despite similar "base-rate" pitch-discrimination sensitivity the SIPI-rate cue dominates also supra-threshold pitch perception.

For inconsistent conditions, contrary to consistent conditions, the effect of ΔAM was medium in size and—contrary to the hypothesis—negative. In addition, this effect seemed to increase with increasing AM depth. However, similar to the consistent conditions, there was a medium positive effect of ΔSIPI which, contrary to the consistent conditions, decreased with increasing AM depth. While these differences between AM depths suggest a more

JASA
https://doi.org/10.1121/10.0019452

prominent role of the AM-rate cue in inconsistent conditions, the negative relation at first sight appears suspicious. To better understand the origin of this effect, we revisited the experimental setup, identified two outlier "single" conditions with ΔAM of 100% and ΔSIPI of 0%, excluded those from the statistical analysis and repeated the rmREGs with the reduced data set. In that analysis the negative effect vanished for the AM depth of 0.1—rendering the inconsistent relation similar to the consistent relation—and declined for the AM depth of 0.5.

These results suggest that the negative effect of ΔAM was a combination of two underlying perceptual effects. First, by contrasting the full data set with the reduced data set, it can be concluded that low rate-discrimination sensitivity in the two outlier conditions was actually a result of ΔSIPI being 0% rather than ΔAM being 100%. Hence, a large part of the negative effect of ΔAM was actually a hidden effect of ΔSIPI. Second, however, the residual (non-significant) effect of ΔAM in the reduced analysis and the increased confidence intervals suggest (1) a residual effect of ΔAM and (2) that this effect was more variable between CI listeners than the ΔSIPI effect.

Taken together, while rate discrimination with consistent cue changes revealed a general perceptual dominance of the SIPI rate, rate discrimination with inconsistent cue changes and an AM depth of 0.5 revealed a larger and more listener-specific contribution of the AM rate to the discrimination judgment. However, also for inconsistent cue changes the SIPI rate remained the dominant rate cue, at least across CI listeners.

McKay and Carlyon (1999) let CI listeners rate the similarity of pairs of low-rate AM pulse trains with both carrier and AM rate varying between 60 and 300 Hz. They found a high degree of listener specificity in the relative contributions of carrier and AM to the composite percept and a systematic increase in the perceptual weight of the AM-rate cue with increasing AM depth. In particular, for AM depths ≥ -4 dB (corresponding to about 0.5), the AM cue seemed to contribute more than the carrier-rate cue. Comparing to our study, we assume that the SIPI-rate cue in our stimuli is comparable to the carrier-rate cue in McKay and Carlyon (1999). Our findings are in line with McKay and Carlyon in that the contribution of the AM-rate cue increased with increasing AM depth. However, contrary to their findings with low-rate carriers, in the present study the AM-rate cue generally seemed to contribute less to the composite percept as compared to the SIPI-rate cue. Thus, for both studies, increasing the sharpness of the AM (which in turn would reduce the "effective" carrier rate as well) appears to increase the salience of the AM-rate cue—at least for temporal-pitch perception. However, even for an AM depth of 1, Goldsworthy *et al.* (2022) found higher pitch resolution (as indicated by lower discrimination thresholds) for pulse-rate as compared to AM-rate cues. They attributed this superiority to higher auditory-nerve synchrony. Similar to this, for ITD sensitivity the SIPI effect was attributed to precise (binaural) spiking (Hancock *et al.*, 2012). Hence, the dominance of the SIPI rate over the AM rate might also be explained in terms of higher SIPI-rate cue accuracy *per se* rather than a larger cognitive weight of the SIPI-rate cue.

B. Effects of average rate and rate limitations

Because our set of conditions involved testing a given Δ at different absolute rates in the range between 63 and 250 Hz, the effects of ΔSIPI and ΔAM were potentially influenced by rate limitations for temporal pitch (cf. Sec. I). To account for such effects, sensitivity parameters for the absolute rates, that is, the average SIPI rate ($\overline{\text{SIPI}}$) and average AM rate ($\overline{\text{AM}}$) were included per condition as predictors in the rmREGs.

Assuming a rate limitation, a negative relation with the d' scores is expected. In consistent conditions, $\overline{\text{AM}}$ indeed tended to show only a small negative relation with the d' scores which was similar across AM depths. In contrast, $\overline{\text{SIPI}}$ showed a medium positive relation, again similar across AM depths. However, there was considerable across-subject variability in this effect. While opposing effects of $\overline{\text{SIPI}}$ and $\overline{\text{AM}}$ might seem counter-intuitive at first, a closer look on the actual ranges of rates involved sheds more light. The range of $\overline{\text{SIPI}}$ (rates between 88 and 213 pps) covered only partially the range of $\overline{\text{AM}}$ (rates between 107 and 213 Hz). Lindenbeck *et al.* (2020) showed decreasing pitch-discrimination performance when lowering the SIPI rate from 125 to 63 pps, while keeping the AM rate at 125 Hz. Such a "low-rate" limitation was also found previously, at least for some listeners (Carlyon *et al.*, 2010; Chatterjee and Oberzut, 2011; Kong *et al.*, 2009; Stahl *et al.*, 2016). In summary, the results suggest that $\overline{\text{AM}}$ was slightly affected by a rate limitation, whereas $\overline{\text{SIPI}}$ was only affected by a "low-rate" limitation.

In inconsistent conditions, there was a weak negative effect of $\overline{\text{AM}}$ at the AM depth of 0.1 but a small positive trend for the AM depth of 0.5. For $\overline{\text{SIPI}}$, the effect was medium, positive, and comparable to consistent conditions for the AM depth of 0.1, however, smaller (but still positive) for the AM depth of 0.5. While the effect of the AM depth *per se* suggests a higher AM-rate cue salience for the larger AM depth, a look at the actual ranges again sheds more light. $\overline{\text{AM}}$ ranged from 107 to 250 Hz (i.e., the upper end of the range was larger compared to consistent conditions) and $\overline{\text{SIPI}}$ ranged from 63 to 177 pps (i.e., the range was shifted towards lower SIPI rates compared to consistent conditions). Consequently, since the rates between 213 and 250 Hz were solely encoded in the AM, the weak AM-rate cue at a depth of 0.1 might have been susceptible to a mild rate limitation kicking in already at rates of approximately 200 pps (Kong *et al.*, 2009, Lindenbeck *et al.*, 2020, Townshend *et al.*, 1987, and Vandali *et al.*, 2013), at least in some CI listeners (Kong and Carlyon, 2010; Kong *et al.*, 2009). Thus, the increase in AM-rate sensitivity with increasing AM depth might be considered as a "release" from that rate limitation due to the increase in the salience of the AM-rate cue. This increased salience might, in combination with the lower $\overline{\text{SIPI}}$ range for inconsistent conditions, explain the effect of

https://doi.org/10.1121/10.0019452 JASA

the AM depth on $\overline{\text{SIPI}}$. In particular, assuming a sufficiently salient AM-rate cue, a range of rates between 177 and 250 Hz solely covered by the AM, and a range of rates between 63 and 107 pps solely covered by the SIPIs, the smaller positive relation might be the result of the cue-specific "sweet spots" forming for the SIPIs at low rates and the AM at high rates. Consequently, this line of argument again supports that, with inconsistent rate cues, both $\overline{\text{SIPI}}$ and $\overline{\text{AM}}$ made relevant contributions to the composite percept. However, the results for $\overline{\text{AM}}$ and $\overline{\text{SIPI}}$ in inconsistent conditions should be interpreted with caution regarding general rate sensitivity for CI listeners.

C. Short inter-pulse intervals for improving both temporal pitch and interaural-time-difference coding?

Insertion of SIPI pulses into high-rate AM pulse trains seems to benefit both temporal-pitch and ITD sensitivity (Lindenbeck *et al.*, 2020; Srinivasan *et al.*, 2020; Srinivasan *et al.*, 2018). However, can the SIPI approach provide both good pitch and good ITD sensitivity, jointly with the same parameter setting?

Temporal-pitch sensitivity improves most when the SIPI rate reflects the AM rate (full-rate SIPI pulses, Lindenbeck *et al.*, 2020). In such a setting, because the AM typically encodes the F0, SIPI rates cover the F0 range, being roughly between 100 and 250 pps (Peterson and Barney, 1952). On the other hand, ITD sensitivity for unmodulated high-rate pulse trains improves most if the SIPI rate does not largely exceed 100 pps (Srinivasan *et al.*, 2018). For high-rate AM pulse trains, ITD-sensitivity improvements were found for a SIPI rate of 63 pps (Srinivasan *et al.*, 2020). A higher SIPI rate can thus be more beneficial for pitch perception than for ITD perception. Taken together, there might be a trade-off between maximizing temporal-pitch and ITD sensitivity, if the SIPI rate is the dominant rate cue (Lindenbeck *et al.*, 2020).

Our results provide evidence that the SIPI rate is indeed the largely dominant rate cue (at least for the AM depths tested, which reflected the typical output of clinical CI processors; see Srinivasan *et al.*, 2017). Consequently, lowering the SIPI rate to a fixed fraction of the AM rate might, while improving ITD sensitivity, have the disadvantage of lowering the perceived pitch height, thus preserving only the relative pitch contour, and of even introducing a second salient pitch cue in case of larger AM depths.

If the efficiency of SIPI pulses turns out to generalize to multi-electrode configurations, then the SIPI approach might be a step towards improving selective hearing. The generalization to multi-electrode stimulation, however, is far from certain: With more than one stimulating electrode, even with CIS-like stimulation, unwanted channel interactions occur (e.g., de Balthasar *et al.*, 2003) with the amount depending on the (carrier) pulse rate. Hence, the amount of interactions might be higher for high-rate AM pulse trains (e.g., McKay and McDermott, 1996) both with and without SIPI pulses as compared to low-rate pulse trains (e.g., Macherey and Carlyon, 2010). If this were the case, multi-electrode pitch sensitivity may suffer more from channel interactions in a high-rate SIPI stimulation paradigm than in a low-rate stimulation paradigm (as provided by stimulation strategies such as FAST or PP; Hossain and Goldsworthy, 2018; Smith *et al.*, 2014; Williges *et al.*, 2018).

V. CONCLUSIONS

In this study, five CI listeners performed pitch discrimination for high-rate AM pulse trains superimposed with SIPI pulses to disentangle the perceptual contributions of AM and SIPI rates to temporal-pitch sensitivity. For consistent changes of AM and SIPI rate, results showed that the SIPI rate dominated pitch perception across the tested AM depths of 0.1 and 0.5. However, in conditions with inconsistent rate-cues changes, the AM seemed to relevantly contribute to the composite percept, in particular, for the larger AM depth. Still, even at the larger AM depth, the SIPI rate appears to dominate pitch perception, despite considerable across-listener variability.

Our results suggest that future strategies aiming to jointly enhance *both* temporal-pitch and ITD sensitivity in CI listeners by implementing the SIPI approach may benefit from listener-specific adjustment of the SIPI rates. Thereby, the potential competition of rate cues will require special consideration, particularly in multi-electrode stimulation.

ACKNOWLEDGMENTS

We thank our CI listeners for their patience and passion. We thank Michael Mihocic for hardware and software support as well as him and Maike Klingel for help with collecting parts of the data. We thank the Institute of Ion Physics and Applied Physics of the Leopold-Franzens-University, Innsbruck, Austria, for providing the equipment for direct electric stimulation. We further thank the associate editor Li Xu and two anonymous reviewers for their fruitful comments on earlier versions of the manuscript. This study was supported by the National Institutes of Health (Grant No. R01 DC 005775). ML was supported by the Austrian Academy of Sciences via a DOC Fellowship (Grant No. A-25606), the Dr. Anton Oelzelt-Newin Fund (Grant No. A-100315), and the Dr. Josef Seegen Endowment Fund (A-100400).

[1]For the sake of comparability, we applied the same criterion also to those three CI listeners that completed the pretest with 100 repetitions. The criterion would otherwise have been $d' = 0.36$.

[2]Depending on the condition, "correct" responses might differ based on which pitch cue was used to make the judgment. Thus, sometimes there was no pre-determinable correct or false and we decided to not provide feedback.

[3]To keep the time gap between carrier pulses and SIPI pulses constant, all SIPI rates had to be integer sub-multiples of the carrier rate. Thus, 91 and 182 Hz are the closest possible realizations of the actual geometric averages (89 and 177 Hz, respectively).

[4]Despite the unbalanced experimental setup, the predictors were sufficiently independent of each other (variance inflation factors ≤ 1.5).

[5]The bootstrap samples were stratified by CI listener and the confidence intervals were corrected for sample bias (cf. Ohtani, 2000).

[6]See supplementary material at https://doi.org/10.1121/10.0019452 for supplementary tables and figures.

JASA
https://doi.org/10.1121/10.0019452

[7]Because the SIPI rate can by definition never plausibly exceed the AM rate, the SIPI approach is inherently unbalanced in that respect. Furthermore, note that a similarly unbalanced design was also employed by McKay and Carlyon (1999).

Arora, K., Dawson, P., Dowell, R., and Vandali, A. (**2009**). "Electrical stimulation rate effects on speech perception in cochlear implants," Int. J. Audiology **48**, 561–567.
Baumann, U., and Nobbe, A. (**2004**). "Pitch ranking with deeply inserted electrode arrays," Ear Hear. **25**, 275–283.
Bland, J. M., and Altman, D. G. (**1994**). "Statistics notes: Correlation, regression, and repeated data," BMJ **308**(6933), 896–896.
Bland, J. M., and Altman, D. G. (**1995**). "Statistics notes: Calculating correlation coefficients with repeated observations: Part 1—Correlation within subjects," BMJ **310**(6977), 446–446.
Bregman, A. S. (**1990**). *Auditory Scene Analysis* (MIT Press, Cambridge, MA).
Canty, A., and Ripley, B. R. (**2022**). "boot: Bootstrap Functions," https://CRAN.R-project.org/package=boot (Last viewed January 16, 2023).
Carlyon, R. P. (**1997**). "The effects of two temporal cues on pitch judgments," J. Acoust. Soc. Am. **102**(2), 1097–1105.
Carlyon, R. P., Deeks, J. M., and McKay, C. M. (**2010**). "The upper limit of temporal pitch for cochlear-implant listeners: Stimulus duration, conditioner pulses, and the number of electrodes stimulated," J. Acoust. Soc. Am. **127**(3), 1469–1478.
Chatterjee, M., and Oberzut, C. (**2011**). "Detection and rate discrimination of amplitude modulation in electrical hearing," J. Acoust. Soc. Am. **130**(3), 1567–1580.
Churchill, T. H., Kan, A., Goupell, M. J., and Litovsky, R. Y. (**2014**). "Spatial hearing benefits demonstrated with presentation of acoustic temporal fine structure cues in bilateral cochlear implant listeners," J. Acoust. Soc. Am. **136**(3), 1246–1256.
Cohen, J. (**1992**). "A power primer," Psychol. Bull. **112**(1), 155–159.
Cousineau, D. (**2005**). "Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method," Quant. Methods Psychol. **1**(1), 42–45.
Croghan, N. B. H., Duran, S. I., and Smith, Z. M. (**2017**). "Re-examining the relationship between number of cochlear implant channels and maximal speech intelligibility," J. Acoust. Soc. Am. **142**(6), EL537–EL543.
Darwin, C. J. (**1997**). "Auditory grouping," Trends Cogn. Sci. **1**(9), 327–333.
de Balthasar, C., Boëx, C., Cosendai, G., Valentini, G., Sigrist, A., and Pelizzone, M. (**2003**). "Channel interactions with high-rate biphasic electrical stimulation in cochlear implant subjects," Hear. Res. **182**(1-2), 77–87.
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.
Gaudrain, E., and Başkent, D. (**2018**). "Discrimination of voice pitch and vocal-tract length in cochlear implant users," Ear Hear. **39**(2), 226–237.
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.
Goldsworthy, R. L., Bissmeyer, S. R. S., and Camarena, A. (**2022**). "Advantages of pulse rate compared to modulation frequency for temporal pitch perception in cochlear implant users," J. Assoc. Res. Otolaryngol. **23**(1), 137–150.
Graves, J. E., and Oxenham, A. J. (**2019**). "Pitch discrimination with mixtures of three concurrent harmonic complexes," J. Acoust. Soc. Am. **145**(4), 2072–2083.
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.
Hochmair, I., Nopp, P., Jolly, C., Schmidt, M., Ser, H. S., Garnham, C., and Anderson, I. (**2006**). "MED-EL cochlear implants: State of the art and a glimpse into the future," Trends Amplif. **10**(4), 201–219.
Hossain, S., and Goldsworthy, R. L. (**2018**). "Factors affecting speech reception in background noise with a vocoder implementation of the FAST algorithm," J. Assoc. Res. Otolaryngol. **19**(4), 467–478.
Ihlefeld, A., Carlyon, R. P., Kan, A., Churchill, T. H., and Litovsky, R. Y. (**2015**). "Limitations on monaural and binaural temporal processing in bilateral cochlear implant listeners," J. Assoc. Res. Otolaryngol. **16**(5), 641–652.
Klein, S. A. (**2001**). "Measuring, estimating, and understanding the psychometric function: A commentary," Percept. Psychophys. **63**(8), 1421–1455.
Kong, Y.-Y., and Carlyon, R. P. (**2010**). "Temporal pitch perception at high rates in cochlear implants," J. Acoust. Soc. Am. **127**(5), 3114–3123.
Kong, Y.-Y., Deeks, J. M., Axon, P. R., and Carlyon, R. P. (**2009**). "Limits of temporal pitch in cochlear implants," J. Acoust. Soc. Am. **125**(3), 1649–1657.
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.
Krumbholz, K., Patterson, R. D., and Pressnitzer, D. (**2000**). "The lower limit of pitch as determined by rate discrimination," J. Acoust. Soc. Am. **108**(3), 1170–1180.
Laback, B., Egger, K., and Majdak, P. (**2015**). "Perception and coding of interaural time differences with bilateral cochlear implants," Hear. Res. **322**, 138–150.
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.
Laback, B., Pok, S.-M., Baumgartner, W.-D., Deutsch, W. A., and Schmid, K. (**2004**). "Sensitivity to interaural level and envelope time differences of two bilateral cochlear implant listeners using clinical sound processors," Ear Hear. **25**(5), 488–500.
Laneau, J., Wouters, J., and Moonen, M. (**2006**). "Improved music perception with explicit pitch coding in cochlear implants," Audiol. Neurotol. **11**(1), 38–52.
Lindenbeck, M. J., Laback, B., Majdak, P., and Srinivasan, S. (**2020**). "Temporal-pitch sensitivity in electric hearing with amplitude modulation and inserted pulses with short inter-pulse intervals," J. Acoust. Soc. Am. **147**(2), 777–793.
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. (**2010**). "Temporal pitch percepts elicited by dual-channel stimulation of a cochlear implant," J. Acoust. Soc. Am. **127**(1), 339–349.
Majdak, P., Laback, B., and Baumgartner, W.-D. (**2006**). "Effects of interaural time differences in fine structure and envelope on lateral discrimination in electric hearing," J. Acoust. Soc. Am. **120**(4), 2190–2201.
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 McDermott, H. J. (**1996**). "The perception of temporal patterns for electrical stimulation presented at one or two intracochlear sites," J. Acoust. Soc. Am. **100**(2), 1081–1092.
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.
McKay, C. M., Remine, M. D., and McDermott, H. J. (**2001**). "Loudness summation for pulsatile electrical stimulation of the cochlea: Effects of rate, electrode separation, level, and mode of stimulation," J. Acoust. Soc. Am. **110**(3), 1514–1524.
Monaghan, J. J. M., and Seeber, B. U. (**2016**). "A method to enhance the use of interaural time differences for cochlear implants in reverberant environments," J. Acoust. Soc. Am. **140**(2), 1116–1129.
Moore, B. C. J., and Gockel, H. E. (**2012**). "Properties of auditory stream formation," Philos. Trans. R. Soc. B **367**(1591), 919–931.
Morey, R. D. (**2008**). "Confidence intervals from normalized data: A correction to Cousineau (2005)," Quant. Methods Psychol. **4**(2), 61–64.
Ohtani, K. (**2000**). "Bootstrapping R^2 and adjusted R^2 in regression analysis," Economic Modell. **17**(4), 473–483.
Peterson, G. E., and Barney, H. L. (**1952**). "Control methods used in a study of the vowels," J. Acoust. Soc. Am. **24**(2), 175–184.
Pressnitzer, D., Patterson, R. D., and Krumbholz, K. (**2001**). "The lower limit of melodic pitch," J. Acoust. Soc. Am. **109**(5), 2074–2084.

https://doi.org/10.1121/10.0019452

JASA

R Core Team (**2022**). *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, Vienna, Austria), https://www.R-project.org/ (Last viewed July 4, 2022).

Riss, D., Hamzavi, J.-S., Blineder, M., Honeder, C., Ehrenreich, I., Kaider, A., Baumgartner, W.-D., Gstoettner, W., and Arnoldner, C. (**2014**). "FS4, FS4-p, and FSP: A 4-month crossover study of 3 fine structure sound-coding strategies," Ear Hear. **35**(6), e272–e281.

Shinn-Cunningham, B. G., and Best, V. (**2008**). "Selective attention in normal and impaired hearing," Trends Amplif. **12**(4), 283–299.

Smith, Z. M., Kan, A., Jones, H. G., Buhr-Lawler, M., Godar, S. P., and Litovsky, R. Y. (**2014**). "Hearing better with interaural time differences and bilateral cochlear implants," J. Acoust. Soc. Am. **135**(4), 2190–2191.

Sourceforge (**2023**). https://sourceforge.net/projects/expsuite/ (Last viewed August 22, 2017).

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 Arnoldner, C. (**2020**). "Improving interaural time difference sensitivity using short inter-pulse intervals with amplitude-modulated pulse trains in bilateral cochlear implants," J. Assoc. Res. Otolaryngol. **21**, 105–120.

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. **19**(3), 301–315.

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

Thakkar, T., Kan, A., Jones, H. G., and Litovsky, R. Y. (**2018**). "Mixed stimulation rates to improve sensitivity of interaural timing differences in bilateral cochlear implant listeners," J. Acoust. Soc. Am. **143**(3), 1428–1440.

Townshend, B., Cotter, N., Van Compernolle, D., and White, R. L. (**1987**). "Pitch perception by cochlear implant subjects," J. Acoust. Soc. Am. **82**(1), 106–115.

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

Vandali, A. E., Sucher, C., Tsang, D. J., McKay, C. M., Chew, J. W. D., and McDermott, H. J. (**2005**). "Pitch ranking ability of cochlear implant recipients: A comparison of sound-processing strategies," J. Acoust. Soc. Am. **117**(5), 3126–3138.

van Hoesel, R. J. M., and Tyler, R. S. (**2003**). "Speech perception, localization, and lateralization with bilateral cochlear implants," J. Acoust. Soc. Am. **113**(3), 1617–1630.

Williges, B., Jürgens, T., Hu, H., and Dietz, M. (**2018**). "Coherent coding of enhanced interaural cues improves sound localization in noise with bilateral cochlear implants," Trends Hear. **22**, 233121651878174.

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.

Supplementary information

I. TABLES

TABLE SI. Goodness of fit for the rmREGs in terms of adjusted R^2 (R_a^2) with 95% confidence intervals (in brackets); consistent (C) and inconsistent (I, both full and reduced [red.] data set) conditions.

Condition	*AM depth*	$F(8, df)$	*df*	*p*	R_a^2
C	0.1	7.13	41	$< .001$	.50 [.28, .58]
C	0.5	6.51	41	$< .001$	.47 [.19, .56]
I, full	0.1	16.99	46	$< .001$	.70 [.49, .78]
I, full	0.5	9.18	46	$< .001$	.55 [.25, .65]
I, red.	0.1	11.02	36	$< .001$	.65 [.40, .71]
I, red.	0.5	2.89	36	.014	.26 [-.03*, .36]

* Note that R_a^2 can be < 0.

TABLE SII. Consistent conditions: Standardized rmREG β coefficients with 95% confidence intervals (in brackets).

AM depth	*Predictor*	β	$t(46)$	*p*	*Fig.*
0.1	ΔAM	-.02 [-.21, .18]	-0.19	.854	5A
0.1	ΔSIPI	.43 [.26, .61]	3.95	$< .001$	6A
0.1	$\overline{\text{AM}}$	-.17 [-.40, .06]	-1.28	.209	S1A
0.1	$\overline{\text{SIPI}}$	.32 [.07, .55]	2.37	.022	S2A
0.5	ΔAM	.00 [-.18, .20]	0.02	.986	5D
0.5	ΔSIPI	.44 [.27, .65]	3.96	$< .001$	6D
0.5	$\overline{\text{AM}}$	-.22 [-.55, .09]	-1.67	.103	S1D
0.5	$\overline{\text{SIPI}}$	.24 [-.03, .58]	1.75	.088	S2D

TABLE SIII. Inconsistent conditions, ***full*** data set: Standardized rmREG β coefficients with 95% confidence intervals (in brackets).

AM depth	*Predictor*	β	$t(46)$	*p*	*Fig.*
0.1	ΔAM	-.31 [-.51, -.15]	-3.81	$< .001$	5B
0.1	ΔSIPI	.48 [.32, .68]	4.96	$< .001$	6B
0.1	$\overline{\text{AM}}$	-.07 [-.29, .14]	-0.67	.504	S1B
0.1	$\overline{\text{SIPI}}$	.35 [.11, .56]	3.64	$< .001$	S2B
0.5	ΔAM	-.47 [-.65, -.27]	4.62	$< .001$	5E
0.5	ΔSIPI	.29 [.05, .49]	2.42	.020	6E
0.5	$\overline{\text{AM}}$	.18 [-.09, .40]	1.29	.205	S1E
0.5	$\overline{\text{SIPI}}$	.17 [-.07, .39]	1.41	.165	S2E

TABLE SIV. Inconsistent conditions, ***reduced*** data set (cf. Tab. SIII).

AM depth	*Predictor*	β	$t(46)$	*p*	*Fig.*
0.1	ΔAM	-.05 [-.16, .23]	0.44	.666	5B
0.1	ΔSIPI	.43 [.24, .72]	3.76	$< .001$	6B
0.1	$\overline{\text{AM}}$	-.13 [-.42, .13]	-0.92	.363	S1B
0.1	$\overline{\text{SIPI}}$	.39 [.11, .66]	3.05	.004	S2B
0.5	ΔAM	-.17 [-.50, .08]	1.12	.270	5E
0.5	ΔSIPI	.28 [-.02, .57]	1.67	.104	6E
0.5	$\overline{\text{AM}}$	.28 [-.16, .61]	1.35	.187	S1E
0.5	$\overline{\text{SIPI}}$	.12 [-.25, .47]	0.65	.523	S2E

II. FIGURES

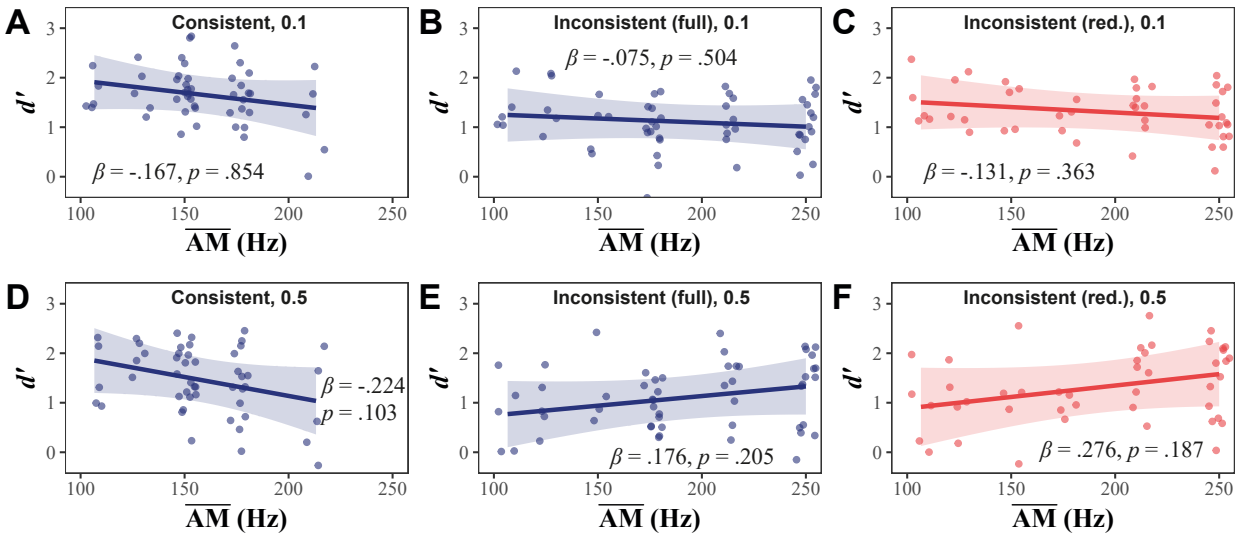


FIG. S1. Relation (i.e., partial correlation) between average AM rate ($\overline{\text{AM}}$) and temporal-pitch sensitivity. Per panel, data set (consistent versus inconsistent [full in blue and red. in red]), AM depth (0.1 and 0.5), standardized regression coefficient β, and statistical significance are denoted.

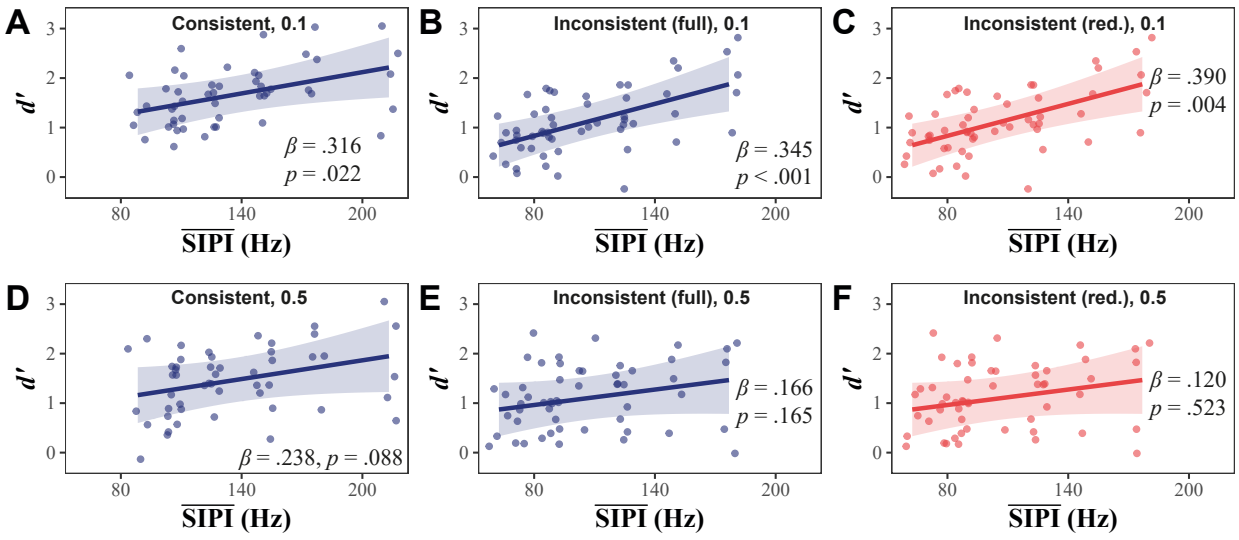


FIG. S2. Relation between average SIPI rate ($\overline{\text{SIPI}}$) and temporal-pitch sensitivity. All other aspects as in Fig. S1.

Channel Interactions with Pulse Rate Cues in Bilateral and Unilateral Dual-Electrode Stimulation

The following manuscript was published in Trends in Hearing:

Lindenbeck, M. J., Majdak, P., & Laback, B. (2024). Effects of Monaural Temporal Electrode Asynchrony and Channel Interactions in Bilateral and Unilateral Cochlear-Implant Stimulation. *Trends in Hearing*, *28*, 23312165241271340. https://doi.org/10.1177/23312165241271340

Author contributions

MJL and BL primarily conceptualized and designed the study. MJL collected the data and performed the statistical analysis. MJL wrote the first draft of the manuscript. All authors contributed to manuscript revision, read and approved the submitted version.

Research Article

Trends in Hearing
Volume 28: 1–18

DOI: 10.1177/23312165241271340
journals.sagepub.com/home/tia
Sage

Effects of Monaural Temporal Electrode Asynchrony and Channel Interactions in Bilateral and Unilateral Cochlear-Implant Stimulation[1]

Martin J. Lindenbeck[1], Piotr Majdak[1] and Bernhard Laback[1]

Abstract

Timing cues such as interaural time differences (ITDs) and temporal pitch are pivotal for sound localization and source segregation, but their perception is degraded in cochlear-implant (CI) listeners as compared to normal-hearing listeners. In multi-electrode stimulation, intra-aural channel interactions between electrodes are assumed to be an important factor limiting access to those cues. The monaural asynchrony of stimulation timing across electrodes is assumed to mediate the amount of these interactions. This study investigated the effect of the monaural temporal electrode asynchrony (mTEA) between two electrodes, applied similarly in both ears, on ITD-based left/right discrimination sensitivity in five CI listeners, using pulse trains with 100 pulses per second and per electrode. Forward-masked spatial tuning curves were measured at both ears to find electrode separations evoking controlled degrees of across-electrode masking. For electrode separations smaller than 3 mm, results showed an effect of mTEA. Patterns were u/v-shaped, consistent with an explanation in terms of the effective pulse rate that appears to be subject to the well-known rate limitation in electric hearing. For separations larger than 7 mm, no mTEA effects were observed. A comparison to monaural rate-pitch discrimination in a separate set of listeners and in a matched setup showed no systematic differences between percepts. Overall, an important role of the mTEA in both binaural and monaural dual-electrode stimulation is consistent with a monaural pulse-rate limitation whose effect is mediated by channel interactions. Future CI stimulation strategies aiming at improved timing-cue encoding should minimize the stimulation delay between nearby electrodes that need to be stimulated successively.

Keywords

interaural time difference, cochlear implant, spatial tuning curve, temporal pitch, discrimination sensitivity, electrode separation

Received 19 October 2023; Revised received 12 April 2024; accepted 24 June 2024

Introduction

Interaural time differences (ITDs) and level differences encode the location of sound sources along the lateral (left-right) dimension and facilitate the perception of a target source in the presence of spatially separated interfering sources (for reviews, see Blauert, 1997; Middlebrooks & Green, 1991). In normal hearing (NH) listeners, ITD is the dominant sound-localization cue at low frequencies (Klingel & Laback, 2022; Macpherson & Middlebrooks, 2002; Wightman & Kistler, 1992). ITD is most saliently conveyed via the rapidly varying temporal fine structure (the carrier signal) in lower frequency channels up to about 1300 Hz (Brughera et al., 2013; Smith et al., 2002; Wightman & Kistler, 1992; Zeng et al., 2004). The salience of ITD in higher frequency channels depends on the peakedness of their temporal envelope (Klein-Hennig et al., 2011; Laback et al., 2011; Macpherson & Middlebrooks, 2002). For signals covering a broader spectral range, NH listeners typically integrate ITD information across frequencies, often resulting in better ITD sensitivity compared to narrowband stimuli (e.g., Buell & Hafter, 1991; Buell & Trahiotis, 1993).

[1]Acoustics Research Institute, Austrian Academy of Sciences, Vienna, Austria

Corresponding author:
Martin J. Lindenbeck, Acoustics Research Institute, Austrian Academy of Sciences, Dominikanerbastei 16, A-1010 Vienna, Austria.
Email: martin.lindenbeck@oeaw.ac.at

The situation is, however, rather different in cochlear-implant (CI) listeners. CIs are auditory prostheses designed to provide speech understanding in profoundly hearing-impaired or deaf people. Modern CIs typically use pulsatile electric stimulation which provides two main advantages. First, the neural response is highly phase-locked to the pulse timing (e.g., Hartmann et al., 1984; Van Den Honert & Stypulkowski, 1987), which is potentially advantageous for encoding timing information at sufficiently low rates (see below). Second, it allows one to avoid simultaneous stimulation at different electrodes and, thus, unwanted electric-field summation that potentially impedes the spectral (across-electrode) stimulus representation (de Balthasar et al., 2003; Eddington et al., 1978; Favre & Pelizzone, 1993). One important aspect in multi-electrode stimulation is the relative timing of pulses *across* electrodes but *within* an ear (e.g., Francart et al., 2015). This is because even with non-simultaneous stimulation of adjacent electrodes—a standard method in clinical CIs—the stimulation pulse pattern received by a certain auditory-nerve sub-population is not simply the pulse pattern of the closest electrode but rather the weighted sum of several nearby electrode signals. This interplay of insufficiently independent CI electrodes is called channel interaction (e.g., de Balthasar et al., 2003).

Such interactions may, in turn, make it difficult for the auditory system to extract the precise timing information that is required to convey ITDs in the range of microseconds. For example, for a certain auditory-nerve population, overlapping stimulation from multiple electrodes may increase the effective rate. If this rate exceeds the above-mentioned rate limitation, ITD coding should worsen. Furthermore, depending on the exact temporal structure of the aggregate pulse pattern, in particular the resulting inter-pulse intervals, temporal effects such as refractoriness or facilitation may incoherently jitter or otherwise degrade the ITD cue even at sufficiently low rates (e.g., Boulet et al., 2016). Hence, appropriately controlling channel interactions may give CI listeners access to lower effective rates and more precise ITD coding.

With a bilateral CI system, every ipsilateral electrode can, in principle, encode ITD when paired with a place-matched contralateral electrode, and the perception of ITDs conveyed by a single interaural electrode pair has been well studied (e.g., Laback et al., 2007; van Hoesel, 2007). The most controlled situation is to stimulate such a single interaural electrode pair using direct stimulation with a research interface and ensuring that the electrodes are interaurally place-matched (i.e., that the stimulated neurons are tonotopically matched and, therefore, compared by binaurally sensitive neurons). Because this configuration avoids channel interactions, it can be used as a baseline for comparison with multi-electrode configurations. Best sensitivity is typically found if the pulse rate is in the order of about 100 pulses per second (pps), and performance is found to decline with increasing pulse rate (an effect referred here to as rate limitation; Laback et al., 2007; Laback et al., 2015; Majdak et al., 2006; van Hoesel, 2007; van Hoesel et al., 2009). While this type of stimulus is sometimes considered as the electrical analog of acoustic fine structure (i.e., "electric fine structure"), there is a remarkable similarity with *envelope* rather than fine-structure ITD sensitivity in NH, both in terms of absolute performance and in terms of the rate limitation (Laback et al., 2015). Sensitivity to envelope ITD in electric hearing, using a high-rate carrier pulse train, shows a similar pattern as for unmodulated pulse trains with matched rates, although absolute performance is lower (Laback et al., 2011) and the decline with rate is sharper (Noel & Eddington, 2013; van Hoesel et al., 2009).

The more practical situation of multi-electrode stimulation has not received as much attention yet. In general, recent findings by Thakkar et al. (2018) and Thakkar et al. (2023) with five-electrode stimulation suggest that coding usable low-rate ITD cues (i.e., 100 pps and, hence, below the rate limitation) on one mid-to-basal electrode yields sensitivity that is significantly better than when stimulating with high rates (i.e., 1000 pps) on all electrodes and comparable to stimulating with low rates on either three or all five electrodes. More specifically, the study by Egger et al. (2016) investigated the tonotopic separation of two electrode pairs and measured pulse-rate ITD thresholds. They stimulated with unmodulated pulse trains at a rate of 100 pps and a delay between these pairs of 5 ms (i.e., half the pulse period) and, therefore, highest potential for detrimental channel interactions due to a doubling of the effective pulse rate. They found a clear effect of electrode separation: In case of wide separations, combined-electrode performance was significantly better than single-electrode performance. In case of small separations, the combined-electrode performance did not differ from single-electrode performance.

With multi-electrode stimulation, also the delay between the pulse trains of two electrodes is an important parameter. In the following, we use the term monaural temporal electrode asynchrony (mTEA) when referring to that delay applied between the pulse trains of the two electrodes in the *same* ear in contrast to the ITD applied *across* ears. Note that the mTEA is a relative measure and is given as a proportion of the pulse period of the single-electrode pulse train. Ihlefeld et al. (2014) measured envelope-ITD sensitivity and used dual-electrode stimulation with 1000-pps pulse trains with 100-Hz sinusoidal amplitude modulation, minimal mTEA, and a relatively wide tonotopic separation. Combined-electrode performance did not differ from the better single-electrode performance. Finally, Francart et al. (2015) measured envelope-ITD sensitivity with 1000-pps pulse trains and a 100-Hz half-wave-rectified sinusoidal AM for both single-electrode and triple-electrode configurations, while varying mTEA from minimal (i.e., almost simultaneous) to maximal (i.e., half the relevant AM period). They found similar sensitivity for triple versus single-electrode stimulation when the mTEA was minimal. With increasing

mTEA, sensitivity worsened, however, only in case of a small separation but not in case of a wider separation. Together, the results of all three studies are consistent with the idea that channel interactions may impede ITD sensitivity in tonotopically proximal multi-electrode stimulation. However, neither of those studies measured the actual degree of channel interaction between the electrodes involved.

In our study, we investigate the hypothesis that channel interactions restrict the access to salient ITD cues with temporally interleaved stimulation. To this end, we studied ITD (i.e., left/right) discrimination sensitivity in bilateral CI listeners being stimulated with two interaurally place-matched electrode pairs. Crucially, each electrode pair used the same stimuli and conveyed the same ITD. Hence, to perform the discrimination task, CI listeners could rely on ITD cues from either electrode pair or a combination of the two. To assess the extent to which channel interactions resulting from dual-electrode stimulation obscure the ITD cues, we varied mTEA and the tonotopic separation, both similarly in the left and the right ear. For large tonotopic separations, we assumed that ITD cues conveyed by each electrode are transmitted through the cochleae independently and are perceptually integrated, hence, in the best case resulting in higher ITD sensitivity for dual- compared to single-electrode stimulation. However, for small tonotopic separations, we assumed that channel interactions impede ITD perception, resulting in similar or even worse performance for dual- compared to single-electrode stimulation.

To achieve a high baseline (i.e., single-electrode) ITD sensitivity, we used pulse trains with 100 pps and per electrode pair (referred to as low-rate pulse trains). To also avoid floor and ceiling effects, we conducted a set of pretests to find, individually per CI listener, one ITD which was used throughout the main experiment. To exclude loudness effects on ITD sensitivity (e.g., Egger et al., 2017) we furthermore loudness-balanced and subsequently image-centered all stimuli.

To alter the *potential* for channel interactions (i.e., the inter-pulse intervals *between* the two electrodes within one ear, cf. Boulet et al., 2016), we systematically varied the mTEA (see the bottom row in Figure 1). The mTEA was varied as a proportion of the pulse period of 10 ms, thus, assuming a periodicity of the mTEA effect. To manipulate the *amount* of channel interactions (i.e., the summation of the two electrodes within a given auditory-nerve neuron), we systematically varied the tonotopic separation between the component electrodes of the dual-electrode combinations, hereafter referred to as C1 and C2, respectively. C1 was fixed in the center of the electrode array while C2 was varied across the array (cf. Figure 2). The selection of tonotopic separations (i.e., C2) was based on a pretest, namely, forward-masked spatial tuning curves (Nelson et al., 2008; Nelson et al., 2011) at C1 in the two ears of each listener.

The fact that the stimuli were completely symmetric across ears highlights that there are in fact two stimulus components contributing to ITD perception, namely the signal at the left ear and the signal at the right ear. One tool to connect the two-ear percept ITD with the perceptual contributions from each of the two ears is *monaural* temporal pitch. As for ITD, temporal pitch is based on the encoding of stimulus timing. In fact, ITD and temporal pitch reveal similar perceptual properties in electric hearing (e.g., Ihlefeld et al., 2015; Kong et al., 2009; Lindenbeck et al., 2020; Lindenbeck et al., 2023). Therefore, similar effects of the parameters mTEA and the tonotopic separation on ITD and temporal pitch may be expected if symmetry across ears is assumed. To that end, our study is complemented by a statistical comparison of the current ITD data with temporal-pitch data collected in unilateral stimulation using matched stimuli (and, hence, rate pitch), mTEA and similar tonotopic separations (Lindenbeck et al., 2024).

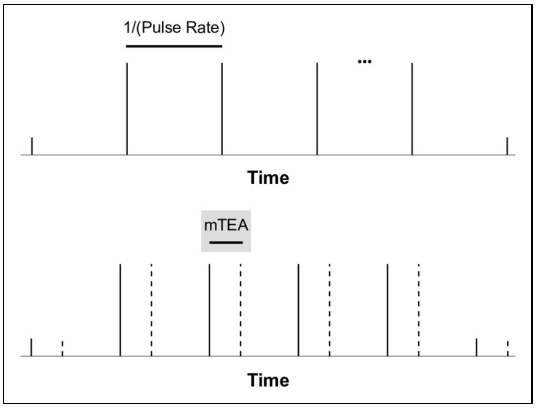


Figure 1. Monaural versions of the unmodulated low-rate stimuli used in the experiment. Top: Single-electrode stimulus. The thick solid horizontal line marks the pulse period (i.e., Pulse Rate^{-1}). Bottom: Dual-electrode stimulus with the monaural temporal electrode asynchrony (mTEA, thick solid horizontal black line in gray box) between basal (solid line) and apical (dashed line) electrode. The mTEA is expressed as a proportion of the pulse period. For readability purposes, negative pulse phases are omitted and only a few pulse periods are shown (indicated by the dots).

Dual-Electrode Interaural-Time-Difference Sensitivity

Methods

Listeners. Five CI listeners (four females) participated, see Table 1 for details. All listeners were post-lingually deafened and used 12-channel implants manufactured by MED-EL Corp. (Innsbruck, Austria) with an electrode spacing of 2.4 mm. In the selection process, we tested listeners having experience in binaural CI testing in our lab and having at least 10 activated electrodes in their clinical fittings. All listeners

Table 1. Details on the CI Listeners.

Listener	Sex	Age (year)	Etiology	Deaf. onset (year)		CI exp. (year)		Array (implant) type	
				L	R	L	R	L	R
CI3	m	37	Meningitis	20	20	17	17	Standard (C40+)	Standard (C40+)
CI12	f	56	VAE; Progressive	32	29	20	22	Standard (C40+)	Standard (C40+)
CI17	f	74	Idiopathic	41	41	5	20	Standard (Synchrony)	Standard (Synchrony)
CI24	f	58	Progressive	40	41	12	11	Standard (C40+)	Standard (C40+)
CI66	f	57	Progressive	36	36	11	10	Standard (Synchrony)	Standard (Concerto)
Mean ± SD	—	56 ± 13	—	34 ± 8	33 ± 9	15 ± 6	16 ± 5	—	—

Note. All listeners used 12-electrode MED-EL CIs in both left (L) and right (R) ears. If the onset of deafness could not be attributed to a specific date (e.g., in case of progressive hearing loss), "deafness onset" may also refer to the onset of profound hearing loss. VAE = vestibular aqueduct syndrome.

were volunteers and were paid an hourly wage for their participation.

In the experiments, we followed the European Charter of Fundamental Rights, worked along the guidelines of "Good Scientific Practice," and fulfilled the ethical principles for research involving human subjects (Helsinki declaration). The research protocol was one of the ethics-approved standard protocols for CI research at the Acoustics Research Institute laboratory. All subjects were adults capable of giving informed consent in writing before the experiment. They were adequately informed of the aims, methods, sources of funding, possible conflicts of interest, institutional affiliations of the researchers, the anticipated benefits and potential risks and discomfort it may entail, post-study support, and any other relevant aspects. They were informed of the right to refuse their participation or to withdraw consent to participate at any time without reprisal. When collecting data, we headed for proportionality and avoided collecting more data than necessary. Confidentiality of collected personal data was maintained by anonymization.

Stimuli and Apparatus. Figure 1 shows monaural and monophasic versions of the biphasic stimuli used in the main experiment. The pulse rate defined the rate at which temporal cues were conveyed. The dual-electrode stimulus (bottom panel in Figure 1) was constructed by presenting the same pulse train at C1 and one of the four C2's. The more apical stimulus (dashed lines) was delayed by the mTEA (i.e., between 0% and 100%). Six of these mTEAs were selected for the main experiment: 4%, 32%, 48%, 52%, 68%, and 96%. These delays were symmetric around 50% and avoided simultaneous stimulation.[2]

Figure 2 shows the four dual-electrode conditions, that is, the four C1–C2 combinations (referred to CCCs), which are named according to the relative position of C2 (i.e., large-apical [L_A], small-apical [S_A], small-basal [S_B], and large-basal [L_B]). In addition to these four CCCs, a fifth "pseudo-dual" electrode condition (called Z for "zero," in blue) was tested in which both C1 and C2 pulse trains were presented at the C1 electrode, simulating a tonotopic separation of zero (i.e., maximum interaction, cf. McKay & McDermott, 1996). The three CCCs S_A, T, and S_B covered electrode separations from 0 to 2.4 mm and, hence, the separations of adjacent electrodes for all three major CI manufacturers in terms of market share, that is, Cochlear Inc. (approximately 0.75 mm), MED-EL (2.4 mm in our cohort), and Advanced Bionics (1.3 mm; Zeng, 2022). The two largest CCCs L_A and L_B were selected in pretests based on spatial tuning curves (see section "Forward-masked spatial tuning curves and C2-electrode selection").

All stimuli were nominally 500 ms long, including linear 100-ms onset and offset ramps. ITDs were applied by delaying the signals of all stimulation electrodes in one ear compared to the other ear. Hence, applying an ITD did not change the mTEA. The pulse duration (including both phases) was 55 μs; phase duration and inter-phase gap depended on the implant type. The stimuli were generated on a personal computer and were sent directly to the implants via the Research Interface Box II (RIB2, Institute of Ion Physics and Applied Physics, Leopold-Franzens-University, Innsbruck, Austria). The RIB2 is bilaterally synced at 600 kHz providing a temporal stimulation precision of 1.67 μs.

Loudness Balancing and Image Centering. After electrode fitting (for details, see section "Electrode pair and ITD selection pretests") and before using a certain stimulus for a sensitivity measurement (i.e., left/right discrimination), in either the pretests or the main experiment described below, the stimulus amplitudes were first monaurally balanced in loudness and afterwards binaurally centered regarding image position. We used the method of adjustment for both loudness balancing and image centering. Provided precautions are taken to avoid starting-level biases (e.g., Van Eeckhoutte et al., 2018) and listeners complying with the task, this method is similarly accurate as adaptive staircase methods (as we have used so far in, e.g., Lindenbeck et al., 2020; Srinivasan et al., 2020) but much faster (about five times in our case, see also Van Eeckhoutte et al., 2018).

For loudness balancing, reference and target stimuli were presented with a 400-ms inter-stimulus break. The reference

was a single-electrode pulse train presented at the C1 electrode. The listeners were asked to indicate whether both stimuli were equally loud and, if not, which one was louder. They were provided with response buttons changing the amplitude by ±2%, ±5%, or ±10% of the dynamic range to iteratively adjust the level of the target stimulus. Per condition, at least two adjustments were performed with the target stimulus level starting above the comfortable level as determined for that electrode in the fittings and at least two adjustments with the target stimulus level starting below the comfortable level. To avoid response bias, the starting level was roved within ±20% of the comfortable level in the single-electrode condition and within ±20% of the balanced single-electrode amplitudes in the dual-electrode condition.

First, single-electrode stimuli were loudness-balanced for all C2 electrodes. Second, using the balanced single-electrode amplitudes, the dual-electrode stimuli with mTEAs of 4%, 32%, and 48% (for more details, see section "Stimuli and apparatus") were loudness-balanced by adjusting their amplitudes jointly on both electrodes by the same proportion of the dynamic range (cf. Macherey & Carlyon, 2010). The amplitudes finally used in the main experiment for mTEAs of 52%, 68%, and 96% (cf. section "Stimuli and apparatus") were copied from the adjustments made for delays of 48%, 32%, and 4%, respectively.

Image centering was conducted similarly to the loudness balancing, except that both ears were stimulated simultaneously, and each trial consisted of a single interval only. Listeners were asked to indicate whether they perceived the stimulus in the center and, if not, whether it was more to the left or right. To adjust the image position, they were provided with six response buttons causing a change in amplitude by ±2%, ±5%, or ±10% of the dynamic range, split 2:1 between the ears in the direction of image movement. Per condition, at least four image centering adjustments were averaged to obtain the final amplitudes used in the sensitivity (i.e., left/right discrimination) tests.

Electrode Pair and ITD Selection Pretests. We conducted a sequence of pretests to find, per CI listener, a suitable set of interaural C1 and C2 electrode pairs as well as a global ITD for all conditions of the main experiment. Listeners were allowed to take breaks at any time.

Electrode Fitting and C1-Electrode Selection. First, we fitted both ears by determining threshold, maximum comfortable level, and comfortable level for all 12 electrodes with unmodulated 500-ms 100-pps pulse trains without ramps using an informal adjustment procedure. Using these fittings, we excluded electrodes inducing uncomfortable sensations and/or having an unexpectedly small dynamic range as compared to the surrounding electrodes.

To allow for as many C2 candidate electrodes as possible, we aimed for a C1 electrode in the center of the array, that is,

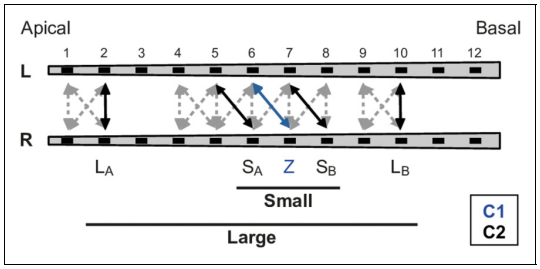


Figure 2. Interaural electrode pair selection. The C1 electrode (Z, in blue) is in the center of the arrangement. Two C2 electrodes are located on the apical side, that is, large-apical (L_A) and small-apical (S_A). Two further C2 electrodes are located on the basal side, that is, small-basal (S_B) and large-basal (L_B). S_A and S_B C2's are located adjacent to C1 (given they provide sufficient masking). L_A and L_B C2's are selected based on forward-masked spatial tuning curves. Arrows with dashed lines indicate interaural electrode pair candidates for the electrode-matching ITD-sensitivity pretest (cf. section "Forward-masked spatial tuning curves and C2-electrode selection").

electrode six or seven (cf. Figure 2). Following evidence that interaural electrode matching is more accurate when based on ITD rather than place pitch (Bernstein et al., 2021; Hu & Dietz, 2015; Long et al., 2003; Poon et al., 2009), we based the interaural electrode-pair selection on ITD. To this end, we tested all four interaural combinations of the two C1 electrode candidates six and seven in an ITD-sensitivity pretest using a two-interval, two-alternative, forced-choice (2I-2AFC) left/right discrimination task and the method of constant stimuli with several pre-selected ITDs. The ITD ranged between 25 and 1200 µs and the individually tested ITDs were selected based on prior knowledge on individual ITD sensitivity from previous studies in our lab. The first stimulus was presented with an ITD of zero and the second stimulus was presented following a 400-ms inter-stimulus break with an ITD pointing either to the left or to the right. Listeners were asked to indicate the perceived side of the second stimulus compared to the first stimulus. Stimuli were presented to both sides 25 times each but in overall randomized order. Response feedback was provided after each trial. The collected percent-correct scores were converted to *d'* scores using the formula $d' = \sqrt{0.5} \cdot [z(P_L) + z(P_R)]$ in which P_L and P_R were the percent-correct scores for left-pointing and right-pointing trials, respectively. This conversion accounted for bias towards either side by averaging over left-pointing and right-pointing trials in the *z*-domain (Klein, 2001). One hundred percent scores were set to 99% to prevent the *d'* from reaching infinity. Hence, the maximum *d'* was 3.3.

The interaural C1 electrode pair yielding the highest *d'* score was selected. All selected C1 electrodes are listed in Table 2.

Table 2. Setup of the Main Experiment as Determined in the Pretests.

	L_A		S_A		Z		S_B		L_B		
Listener	L	R	L	R	L	R	L	R	L	R	ITD (μs)
CI3	3	2	5	6	6	7	8	8	11	11	150
CI12	2	2	5	5	6	6	7	7	9	9	200
CI17	3	3	5	6	7	7	8	8	9	9	1200
CI24	1	2	5	5	6	6	7	8	9	9	200
CI66	1	2	5	5	6	6	7	7	10	9	1000
El. #: mean	2.0	2.2	5.0	5.4	6.2	6.4	7.4	7.6	9.6	9.4	—
El. #: SD	1.0	0.4	0.0	0.5	0.4	0.5	0.5	0.5	0.9	0.9	—
Distance *re* Z (mm): mean	−10.1	−10.1	−2.9	−2.4	—	—	2.9	2.9	8.2	7.2	—
Distance *re* Z (mm): SD	2.0	1.1	1.1	0.0	—	—	1.1	1.1	2.7	1.7	—

Note. Per CI listener: the C1 (Z) and four C2 electrodes. Statistics for electrode number (El. #) and distance (in mm) from the C1 electrode Z, separately for left (L) and right (R) ears. Electrodes are numbered from apex to base. C1–C2 combination (CCC): L_A (large-apical), S_A (small-apical) S_B (small-basal) L_B (large-basal).

Forward-Masked Spatial Tuning Curves and C2-Electrode Selection. The C2 electrode pair candidates and, exemplary, the finally selected electrode pair configurations are shown in Figure 2. We aimed at selecting two types of C2 electrodes: C2's with a small tonotopic separation from C1 to induce maximum channel interactions and C2's with a large tonotopic separation from C1 to induce as little channel interactions as possible. As a marker of channel-interaction potential (i.e., across-electrode masking), we measured listener-specific forward-masked spatial tuning curves using the paradigm described in Nelson et al. (2008, 2011).[3]

Before measuring the tuning curves, we fitted both ears for the tuning-curve stimuli using the same procedure as described in section "Electrode fitting and C1-electrode selection." As masker stimulus, we used an unmodulated 160-ms 1000-pps pulse train without ramps. As target stimulus, we used an unmodulated 10-ms 1000-pps pulse train without ramps on the C1 electrode only. The masker was presented at all available electrodes, whereas the target was always presented at the C1 electrode with a masker-target offset-to-offset delay of 20 ms (cf. Nelson et al., 2008).

The tuning curves were measured separately per ear. The target was presented at the level of 16.5% ± 6.5% (range 7.5%–25.0%) of the target's dynamic range on both ears. The target level was adjusted individually per ear to minimize floor and ceiling effects due to task difficulty but remained fixed during the measurement of the tuning curve. Target levels in this range should not have had an effect on the tuning curve's shape (Nelson et al., 2008). The masker level was varied and the level at threshold was determined in an adaptive 2-up 1-down staircase procedure (Leek, 2001; Levitt, 1971) converging at 71% of the targets being correctly identified. The task was a three-interval three-alternative forced-choice (3I-3AFC) task. Each interval included the masker. In one randomly chosen interval, the masker was followed by the target. The listeners were instructed to indicate the interval that included the target. The initial masker level was 10% of the masker dynamic range. The initial step size was 5% of the masker dynamic range. It was reduced by a factor of 0.7 after each reversal until it reached the minimum step size of 1% of the masker dynamic range. A converging staircase was terminated after 12 reversals. A diverging staircase either exceeded the maximum comfortable level (i.e., it was aborted before stimulating at too loud levels) or dropped below the absolute threshold of the masker electrode. The last eight reversals of a converged staircase were averaged to determine the staircase threshold. The final masking threshold was determined as the average of three staircase thresholds.

After having measured the tuning curves, the small-separation C2 electrode pair candidates (one or two neighboring electrodes of C1 on either side) were evaluated for sufficient masking on both ears. In addition to the presence of masking, the small-separation pairs were required to have ITD sensitivity similar to that of the C1 electrode pair. To find suitable pairs, we conducted another electrode-matching ITD-sensitivity pretest as described in section "Electrode fitting and C1-electrode selection." The finally selected small-separation pairs are summarized in Table 2.

The large-separation C2 electrode pairs were selected based on three criteria. The first criterion was the presence of nearly no masking of C1 at both ears, as indicated by diverging tuning-curve staircases (i.e., masker level exceeding the maximum comfortable level) or, at least, largest masked thresholds on the respective side of C1. The second criterion was the tonotopic separation from C1: We estimated the extent of channel interactions by selecting the closest possible non-masking C2. The third criterion was—similarly to the small-separation C2's—ITD sensitivity comparable to that of the C1 electrode pair. That criterion was again assessed with an electrode-matching ITD-sensitivity pretest similar to that used for the small-separation C2's. For this test, candidate electrode pairs were selected based

on the first two criteria and typically were the closest and second closest possible electrodes. The finally selected large-separation pairs are summarized in Table 2.

Main Experiment: Listener-Specific Interaural Time Difference. Table 2 summarizes the listener-specific setup obtained for the main experiment. The small tonotopic separation was on average and across ears less than 3 mm (i.e., roughly one electrode). The wide separation was on average about 7–8 mm on the basal side and about 10 mm on the apical side.

To maximize the number of testable mTEA and CCC conditions in the main experiment, and to minimize floor and ceiling effects, the listeners were tested at only one listener-specific ITD (Ihlefeld et al., 2015; Kong et al., 2009). The listener-specific ITD used in the main experiment was selected based on the results of the electrode-matching ITD-sensitivity pretests (cf. section "Electrode fitting and C1-electrode selection" and "Forward-masked spatial tuning curves and C2-electrode selection") and aimed at a homogeneous sensitivity across electrode pairs, optimally, slightly above threshold (i.e., a *d'* at or slightly above 1 across electrode pairs). The ITDs selected for individual listeners are summarized in Table 2. Single-electrode ITD sensitivity determined in the pretests was heterogeneous across listeners. Hence, three listeners were tested with small ITDs ($\leq$ 200 μs), and two listeners were tested with large ITDs ($\geq$ 1000 μs).

In the main experiment, the left/right discrimination task was identical to that used in the electrode-matching ITD-sensitivity pretests (see section "Electrode fitting and C1-electrode selection"). Per-trial feedback was provided. Scores were again quantified in terms of *d'*. To keep place cues constant within a block, the CCCs were tested block-wise. The CCC block order was balanced across listeners but different for each individual listener to (pseudo-)randomize block-order effects. Within each CCC block, the order of mTEAs was randomized. To disentangle interaction effects arising from dual-electrode stimulation from electrode-specific baseline ITD sensitivity, single-electrode conditions, corresponding to C1 and C2 of all dual-electrode conditions, were also measured before dual-electrode conditions within each CCC block. Finally, the entire setup described above (i.e., both single- and dual-electrode conditions) was tested twice, each time with 50 repetitions per condition, resulting in 3500 trials in total per listener.

Statistical Analysis. The main goal of our statistical analysis was to dissociate both the presence and the absence of an effect from measurement noise. In classical terms, we aimed to collect evidence for both the null hypothesis (H0, effect absent) and the alternative hypothesis (H1, effect present). However, the most used statistical method of separating systematic effects from measurement noise using null-hypothesis significance testing (NHST) based on *p* values only insufficiently tests the H0 and has been extensively criticized (e.g., Amrhein et al., 2019; Cohen, 1994; Rouder et al., 2016). Among others, two recently proposed alternatives are reporting confidence intervals related to the *p* values (e.g., Cumming, 2014) and Bayesian hypothesis testing (e.g., Wagenmakers et al., 2017). While confidence intervals (e.g., alongside effect sizes) provide both an assessment of statistical power and a measure of practical relevance of a statistical effect, Bayesian hypothesis testing is conceptually different and provides information about the *relative* evidence for H1 over H0 and *vice versa*. Thus, we report our statistical results as Bayes-NHST hybrids (cf. Dienes & Mclatchie, 2018; Keysers et al., 2020). This allows us to collect H0 evidence but still retain a fixed Type I error rate and provide estimates of statistical power.

Bayesian Hypothesis Testing. We used repeated-measures analyses of variance (rmANOVAs) for statistical analysis. In the Bayesian implementation (cf. Keysers et al., 2020; van den Bergh et al., 2020) the Bayes factor (BF; e.g., Kass & Raftery, 1995) is used to distinguish between "evidence of presence" (i.e., diagnostic data supporting H1 over H0), "evidence of absence" (i.e., diagnostic data supporting H0 over H1), and "absence of evidence" (i.e., non-diagnostic data that support neither hypothesis). We considered BFs > 3 as evidence for H1 over H0 ("evidence of presence") and used the inverse threshold (i.e., BF < 0.33) to establish the evidence of H0 over H1 ("evidence of absence"). BFs between 0.33 and 3 were considered inconclusive ("absence of evidence").

Bayesian statistical analyses were conducted with JASP v0.17.1 (JASP Team, 2023). In particular, we used Bayesian rmANOVAs (Rouder et al., 2012, 2016, 2017; van den Bergh et al., 2020, 2022; Wetzels et al., 2012) in the default JASP prior configuration. The output is the so-called BF_{incl} that quantifies evidence for including an effect in the model (H1) over excluding it (H0).

Null-Hypothesis Significance Testing and Confidence Intervals. To provide effect size estimates and well-known Type I error control, we also conducted "classical" NHST rmANOVAs using MATLAB R2022b (Mathworks, Inc). Both Bayesian and NHST results are always reported together. The significance level was always 5%. To prevent inflation of Type I errors, we used the Huynh–Feldt correction (Huynh & Feldt, 1970) of *p* values and *F*-test degrees of freedom (cf. Oberfeld & Franke, 2013).

As effect-size metric, we used the generalized η^2 (η^2_G) (Bakeman, 2005; Lakens, 2013; Olejnik & Algina, 2003). It is comparable across both within-subjects and between-subjects designs and classifies effect sizes as either small ($\eta^2_G \geq .02$), medium ($\eta^2_G \geq .13$), or large ($\eta^2_G \geq .26$) (Bakeman, 2005). The so-called exact analytical 90% "within-subjects" confidence intervals (i.e., the 5%–95% range of η^2_G, the *F*-test is one-sided) are reported in brackets

following the η^2_G estimate. Note that the confidence interval of a statistically significant effect does not contain zero.

Results and Discussion

Single-Electrode Stimulation. Figure 3 shows the listener-averaged *d'* scores for the single-electrode conditions, denoted with an "S". The individual listener scores are shown in Figure 4. Across listeners, basal single-electrode ITD sensitivity might appear to be higher than apical sensitivity but note the large variance (particularly due to CI3 and CI66) for the basal-most (L_B) electrode. However, a one-way rmANOVA on the *d'* scores with the factor electrode revealed evidence for the absence of an effect of electrode ($BF_{incl} = 0.28$, $F[2.1, 8.5] = 0.52$, $p = .621$, $\eta^2_G = .09$ [.00, .28]).

Dual-Electrode Stimulation. The listener-averaged *d'* scores for the dual-electrode conditions as a function of the mTEA (in %) are shown in Figure 3. The individual listener scores are shown in Figure 4. The CCCs are shown in separate panels categorized by their tonotopic separation, that is, zero (left), small (about one electrode, middle), and large (three to four electrodes, right). Conditions with apical and basal C2's are denoted by empty and filled circles, respectively.

To jointly assess the effect of mTEA depending on the tonotopic separation statistically, we conducted a two-way rmANOVA. The factor CCC encodes tonotopic separation, that is, Z for zero separation, S_A and S_B for small separation, and L_A and L_B for large separation, respectively (cf. Figure 2). We found evidence for an mTEA effect ($BF_{incl} = 16.99$, $F[4.5, 17.8] = 8.71$, $p < .001$, $\eta^2_G = .14$ [.07, .19]) that depended on the CCC (mTEA × CCC interaction; $BF_{incl} = 15.17$, $F[20.0, 80.0] = 2.15$, $p = .009$, $\eta^2_G = .12$ [.02, .12]). Furthermore, we found evidence for the absence of a CCC effect ($BF_{incl} = 0.23$, $F[4.0, 16.0] = 0.56$, $p = .693$, $\eta^2_G = .04$ [.00, .08]).

To assess the differences in the effect of the mTEA for the CCCs tested, we conducted five additional one-way rmANOVAs, one per CCC and each including only the factor mTEA. The results are summarized in Table S1 and visualized in Figure 5.

For the zero-separation CCC (Z, left panel), we found evidence for a large mTEA effect. The drop in (pseudo-)dual-electrode sensitivity *re* the single-electrode sensitivity was smallest (about −0.2) for short delays of 4% and 96%, respectively, and considerably larger (about −0.8) for all other delays with a maximum drop for 48%. This effect was largely consistent across listeners with the notable exception of CI24 who showed no mTEA effect.

For small-separation CCCs (middle panel) and the apical condition S_A (empty markers), there was inconclusive evidence only. When compared to the zero-separation CCC, the drop in sensitivity for delays between 32% and 68% was less pronounced whereas it was more pronounced for 96%. The effect differed considerably between listeners, with CI66 showing no mTEA effect and both CI12 and CI3 showing only little effect. In contrast, we found evidence for a large mTEA effect for the basal CCC (S_B, filled

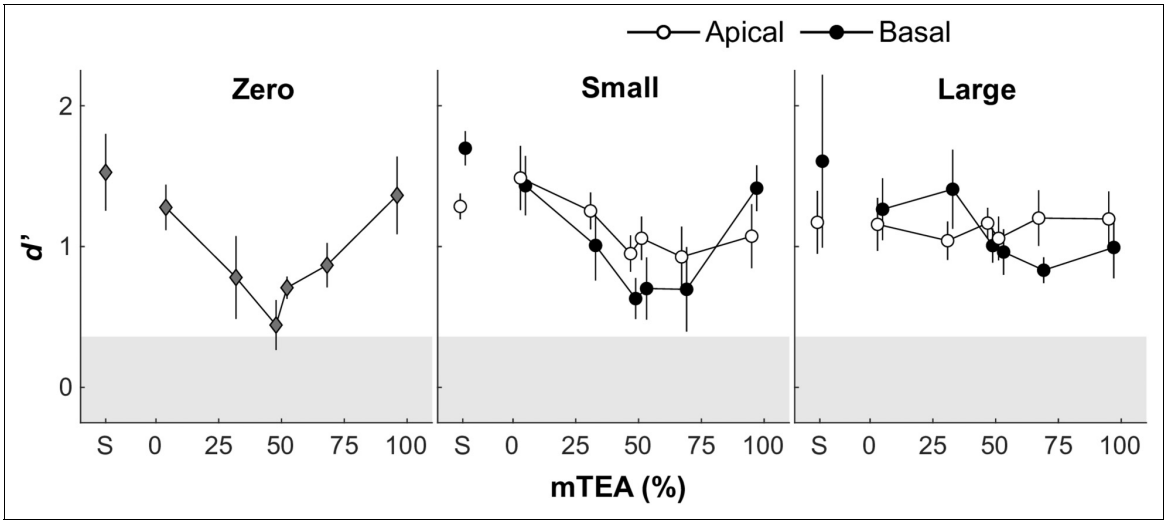


Figure 3. ITD-based left/right discrimination sensitivity *d'* as a function of mTEA (expressed as a proportion of the pulse period) averaged across CI listeners. Panels distinguish tonotopic separations and C1–C2 combinations (CCCs) are named after the C2 electrode involved (cf. Figure 2). For small and large separations, apical CCCs (S_A and L_A, respectively) are shown with empty markers and basal CCCs (S_B and L_B, respectively) are shown with filled markers. For each CCC, single-electrode scores for the C2 electrode are denoted with an "S." The grey area indicates chance performance ($|d'| \leq .36$). Error bars show normalized standard errors (Cousineau, 2005; Morey, 2008).

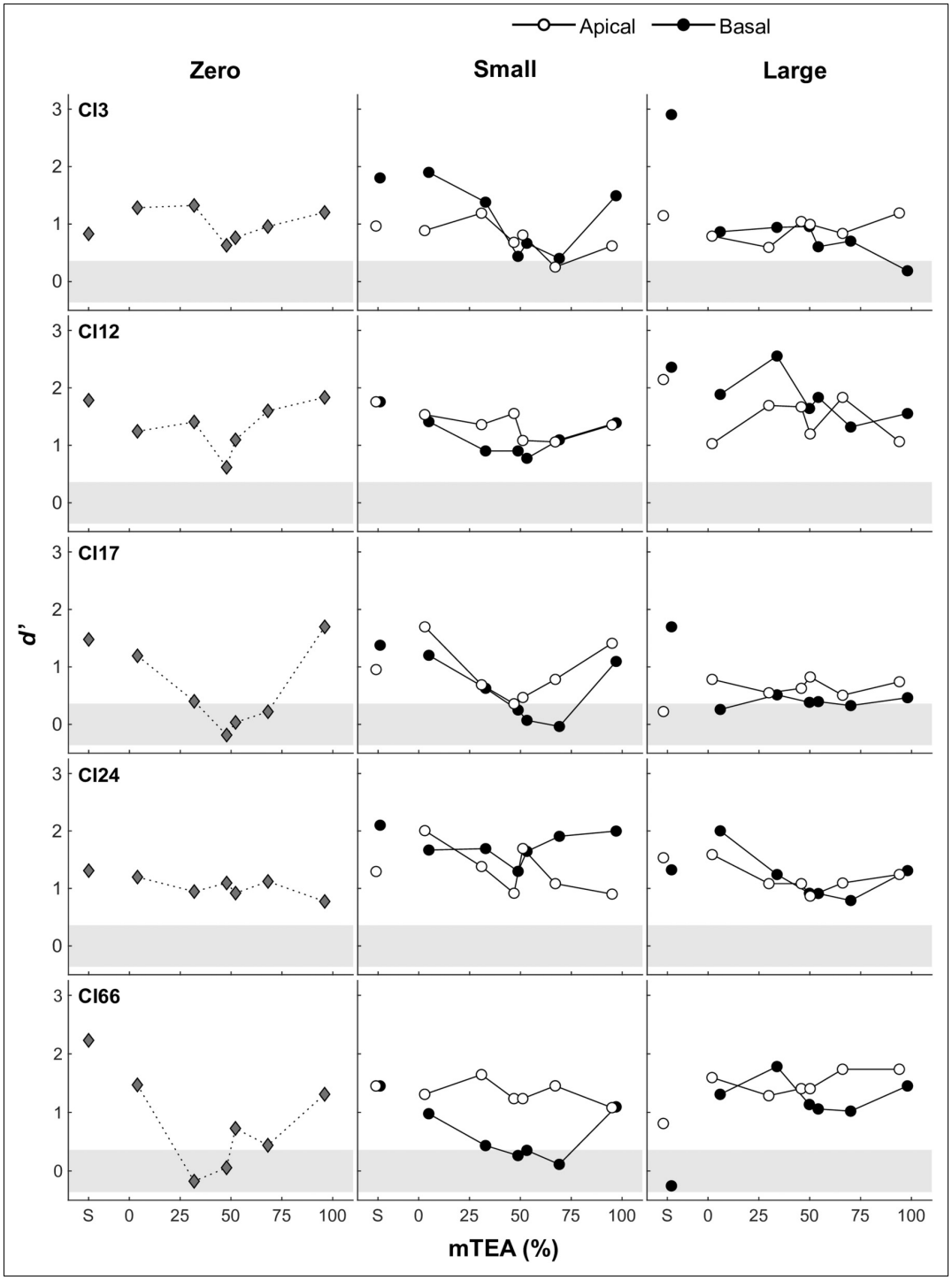


Figure 4. ITD-based left/right discrimination sensitivity *d'* as a function of mTEA in separate rows per CI listener. All other aspects as in Figure 3.

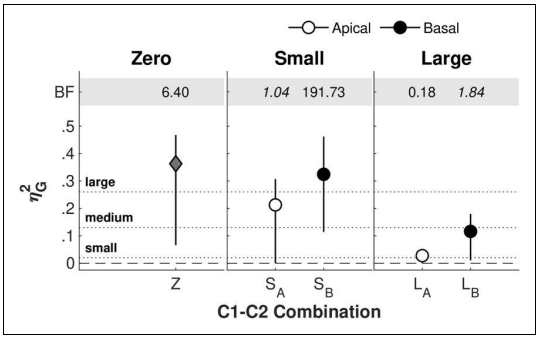


Figure 5. Effect-size estimates η^2_G alongside 90% confidence intervals for several tonotopic separations (panels) and corresponding C1–C2 combinations (CCCs, abscissa). Zero separation: Z CCC. Small separation: apical (S_A, empty marker) and basal (S_B, filled marker) CCC. Large separation: apical (L_A, empty marker) and basal (L_B, filled marker) CCC. The dashed line indicates an η^2_G of zero; effects whose confidence intervals do not touch the dashed line are significant with $p < .05$. Dotted lines indicate threshold η^2_G for small, medium, and large effects, respectively. Bayes factors (BFs) in the grey ribbon indicate evidence for the presence of an mTEA effect (BF > 3) *re* evidence of its absence (BF < 0.33), given our data (BF_{incl}, cf. section "Statistical analysis"). BFs in italics indicate inconclusive results (i.e., absence of evidence). Color and marker-shape coding is consistent with Figure 3.

markers). Across listeners, the S_B mTEA effect was more homogeneous than the S_A mTEA effect, with CI24 as an exception showing only little S_B mTEA effect.[4] The Z and S_B mTEA effects closely resemble each other.

For large-separation CCCs (right panel) and the apical condition L_A (empty markers), we found evidence (homogeneous across listeners) for the absence of an mTEA effect. In contrast, there was inconclusive evidence for the basal condition L_B (filled markers). This might have been due to non-negligible effects of at least some mTEAs for CI12 and CI24.

In summary, our hypothesis on the mTEA effect for the zero and small separations was clearly confirmed for two of the three relevant conditions. For condition S_A, our results are inconclusive, showing neither clear evidence for absence nor for presence of an mTEA effect. For the large-separation CCCs, we hypothesized the absence of an mTEA effect and we found that evidence indeed, however, in the L_A condition only. In the L_B condition, our results are inconclusive.

Comparison Between Dual-Electrode Interaural-Time-Difference and Rate-Pitch Sensitivity

In our behavioral ITD experiment, both parameters mTEA and CCC were *monaural*, which, in case of bilateral stimulation and ITD perception, affected peripheral processing in both ears (cf. Ihlefeld et al., 2015). To provide salient ITD cues, two decent monaural electrode–neuron interfaces are necessary, as compared to only one interface for monaural cues such as rate-based pitch. Hence, the effects of mTEA and CCC might manifest similarly in monaural rate-pitch and ITD perception.

While both percepts are based on similar electrode–neuron interfaces (apart from, e.g., insertion depth and electrode-to-modiolus distance), properties of upstream processing can be percept-specific, for example, pitch-selective neurons in auditory cortex (e.g., Bendor & Wang, 2005) or ITD-sensitive neurons in the brainstem (e.g., Grothe et al., 2010). A study by Ihlefeld et al. (2015) compared rate-limitation effects in monaural rate-pitch versus ITD sensitivity and found ITD sensitivity to correlate significantly with worse-ear rate-pitch sensitivity only. This suggests that a major determinant of ITD sensitivity is the poorer ear which then acts as a monaural bottleneck. Differences between ears in terms of speech recognition performance have further been shown to be related to binaural benefit (Yoon et al., 2011). Still, the rather small explained variance of the correlation in Ihlefeld et al. (2015) suggests additional and possibly percept-specific effects. Note, however, that this correlation was probably underestimated, because it was purely within-subjects and, hence, may have partialed out listener-specific *sensory* effects common to both tasks.

To further isolate the mechanism underlying the mTEA effect for ITD perception, we compared our ITD-based data with data collected for monaural rate-pitch perception in a matched experimental setup but a different cohort of CI listeners (reported in Lindenbeck et al., 2024). The comparison aimed to minimize the confounding effects of differences between the two CI cohorts, particularly the influence of baseline single-electrode sensitivity, task-specific sensitivity, listener-specific amounts of channel interactions (for large tonotopic separations), and both floor and ceiling effects in general. To this end, we individualized important aspects of the experimental setup (i.e., ITD, rate difference, and tonotopic separation) to test the CI listeners at roughly similar points on their individual psychometric functions (corresponding to perceived difficulty) rather than with similar physical properties.

Methods

Rate-Pitch Data Set. The rate-pitch data set from Lindenbeck et al. (2024) was matched as closely as possible to our ITD data set. In that rate-pitch study, five CI listeners (55 ± 9 years old, four females) were tested monaurally. Three listeners were bilaterally implanted and tested with their preferred ear. The age at onset of deafness was 45 ± 11 years and the CI experience was 7 ± 6 years. All listeners had 12-channel MED-EL implants. Two listeners had electrode arrays with 2.4 mm spacing and three listeners had electrode arrays

with 2.1 mm spacing. Note that CI24 participated in both the rate pitch and the ITD experiment.

The C1 electrode was electrode six, thus, located in the center of the electrode array. Small-separation C2's S_A and S_B were the neighbors of C1 and, hence, spaced on average about 2.2 mm apart. Large-separation C2's L_A and L_B were selected based on forward-masked tuning curves (cf. section "Forward-masked spatial tuning curves and C2-electrode selection") and were located on average about 9 mm apically and about 7 mm basally of C1, respectively.

In the experiment, we studied the sensitivity to rate difference (Weber fraction in %) which was fixed for each CI listener, amounting to 15% ± 7%. Analogous to the current ITD data, the rate difference was determined individually per CI listener in single-electrode pretests using unmodulated low-rate stimuli on all five electrodes selected. The tested geometric average (i.e., nominal) rate (cf. Kreft et al., 2010; Lindenbeck et al., 2020) was as close as possible to 100 pps (for more details, see Lindenbeck et al., 2024).

Data Preparation. To remove residual differences in baseline single-electrode sensitivity, we converted the absolute dual-electrode *d'* scores into relative dual-electrode *d'* scores (hereafter referred to as Δ*d'*) by subtracting the Pythagorean average $\overline{d'_j}$ (cf. Green & Swets, 1966) of the C1-electrode-only score $d'_{\text{single, Z}}$ and the C2-electrode-only score $d'_{\text{single, j}}$ (for C2 *j*) from the dual-electrode *d'* of the corresponding CCC.[5] Hence, for mTEA *i* and C2 *j*, the score was calculated as $\Delta d'_{ij} = d'_{ij} - \overline{d'_j}$ with $\overline{d'_j} = \text{sign}(X) \cdot \sqrt{|X|}$ and $X = 0.5 \cdot (\text{sign}(d'_{\text{single, Z}}) \cdot d'^2_{\text{single, Z}} + \text{sign}(d'_{\text{single,j}}) \cdot d'^2_{\text{single,j}})$. In doing so, Δ*d'* reflects dual-electrode ITD sensitivity *re* the prediction of optimal integration of information assuming independent peripheral noise sources for the C1 and C2 electrodes. The Δ*d'* metric furthermore corrects for loudness-balancing performed for dual-electrode conditions with the comfortably loud single-electrode C1-electrode condition via the factor of 0.5 in *X* (for details see section "Loudness balancing and image centering").[6] Hence, lower per-electrode stimulus levels in dual-electrode conditions due to loudness matching of dual-electrode and single-electrode conditions should reduce the amount of information available to the auditory system in dual-electrode stimulation even under ideal conditions (cf. Egger et al., 2016). In fact, considering the zero-separation CCC as a simplified reference, the integrated *d'* scores should be identical to the single-electrode *d'* of the C1 electrode. To this end, the independent-noise integration $d' = \sqrt{d'_Z{}^2 + d'^2_Z} = \sqrt{2} \cdot d'^2_Z$ (Pythagorean sum) has to be multiplied by $\sqrt{0.5}$.[7] Note that this prediction ignores any limiting factors such as channel interactions or pulse rate limitation, thus any negative deviation from the prediction is an indirect indicator of the involvement of such factors.

Both experiments involved only a single value of the FD or ITD per CI listener, selected based on pretests conducted with unmodulated single-electrode low-rate pulse train stimulation on all five electrodes (pairs, in case of ITD). The FDs and ITDs were selected to "calibrate" the setup such that the absolute dual-electrode *d'* scores showed as little floor and ceiling effects as possible (cf. section "Pretests"). Perfectly calibrated, all listeners would have—on average—performed similarly, assuming no other differences between percepts. Because average sensitivity was better than chance level and did not reach the ceiling in all conditions irrespective of the percept, we consider the Δ*d'*-based analysis to yield a valid assessment of between-percept differences.

Statistical Analysis. To search for differences between percepts (i.e., ITD and rate pitch) we conducted an analysis of variance with both between-subjects and within-subjects effects (mixed ANOVA). The between-subjects factor was the Percept (ITD and rate pitch). Within-subjects factors were the mTEA (4%, 32%, 48%, 52%, 68%, and 96%) and the CCC (Z, S_A, S_B, L_A, and L_B). All other aspects were as described in sections "Bayesian hypothesis testing" and "Null-hypothesis significance testing and confidence intervals."

Results and Discussion. Figure 6 shows the Δ*d'* scores as a function of the mTEA in separate panels for each CCC and with the percept as parameter (ITD with empty markers; pitch with filled markers). Note that negative Δ*d'* scores indicate a deterioration and positive Δ*d'* scores indicate an improvement relative to the prediction of optimal integration across electrodes. The error bars show the 95% credible intervals as provided by the "descriptives plots" option of the Bayesian rmANOVA routine in JASP.

When visually comparing between the two percepts, for the zero-separation CCC (Z, leftmost panel), the rate-pitch mTEA effect had a u-shape similar to the ITD mTEA effect. In fact, the u-shape for pitch was even a bit broader than that for ITD, suggesting that the mTEA effect might have been larger for pitch than for ITD. For small tonotopic separations, in particular for S_A (second-to-left panel), there was a clear u/v-shaped mTEA effect for pitch, but little to no mTEA effect for ITD. At least for some mTEAs, the differences between ITD and pitch were larger than the credible intervals. For S_B (center panel), the mTEA effect was similar for ITD and pitch. For large tonotopic separations, in particular L_A (second-to-right panel), there was no mTEA effect, neither for ITD nor pitch, even though the ITD Δ*d'* scores tended to be lower than the pitch Δ*d'* scores. For L_B (rightmost panel), there was no mTEA effect for pitch, however, for ITD, we observed a drop in Δ*d'* from 32% to 48%.

To assess statistical differences between ITD and pitch discrimination sensitivity, we first analyzed all interaction terms of the mixed ANOVA that included the factor Percept. We found evidence for the absence of an effect of Percept on the mTEA effect (Percept × mTEA interaction; $BF_{\text{incl}} = 0.16$, $F[2.2, 17.4] = 0.87$, $p = .443$, $\eta^2_G = .01$ [.00, .04]). The evidence for the other interactions was

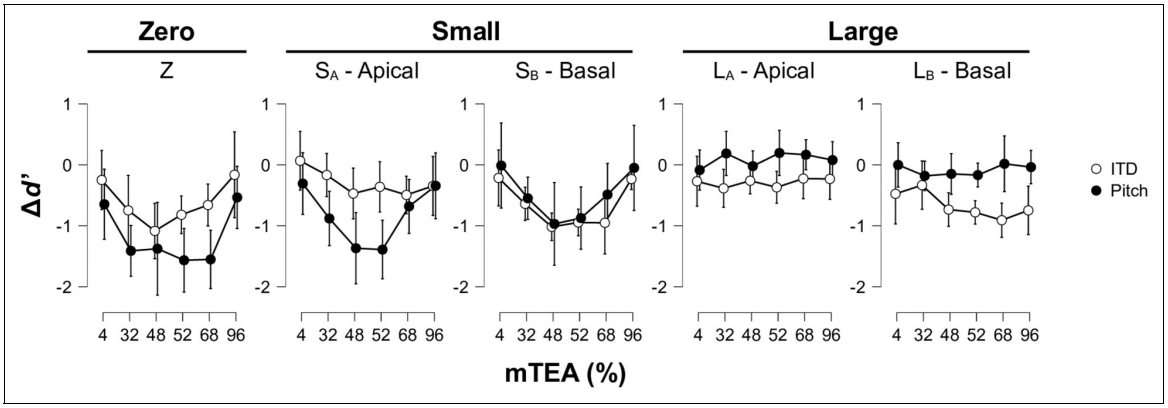


Figure 6. Comparison of ITD-based left/right (empty markers) and rate-pitch (filled markers) discrimination sensitivity in terms of $\Delta d'$ (for more details, see section "Data preparation") as a function of mTEA (cf. Figure 1) and in separate panels per C1–C2 combination (CCC, cf. Figure 2). Error bars show Bayesian 95% credible intervals (JASP "descriptives plots" output; JASP Team, 2023).

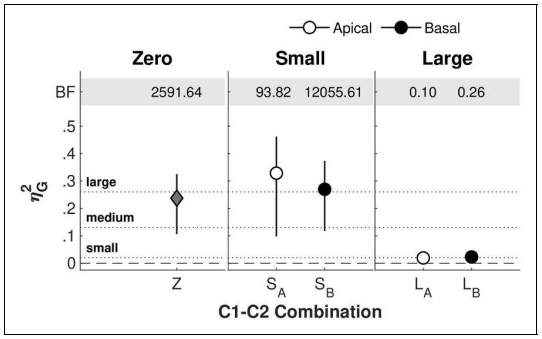


Figure 7. Effect of mTEA (i.e., η^2_G alongside 90% confidence intervals) *across* percepts for several tonotopic separations (panels) and corresponding C1–C2 combinations (abscissa). All other aspects as in Figure 5.

inconclusive (Percept × mTEA × CCC interaction: $BF_{incl} = 2.08$, $F[11.5, 92.3] = 1.72$, $p = .078$, $\eta^2_G = .04$ [.00, .05]; Percept × CCC interaction: $BF_{incl} = 2.59$, $F[2.8, 22.4] = 3.18$, $p = .046$, $\eta^2_G = .15$ [.00, .29]). The main effect of Percept, although inconclusive, was negligible in size ($BF_{incl} = 0.46$, $F[1.0, 8.0] = 0.00$, $p = .967$, $\eta^2_G = .00$ [.00, .00]).

Subsequently, after having established that the mixed ANOVA showed no clear evidence for effects of or interactions with Percept, we also analyzed the effects when pooling the data across percepts. We found evidence that the mTEA effect differed between CCCs (mTEA × CCC interaction; $BF_{incl} = 41231.17$, $F[11.5, 92.3] = 4.09$, $p < .001$, $\eta^2_G = .09$ [.03, .11]). Furthermore, we found evidence for a medium mTEA effect across CCCs ($BF_{incl} = 16321.46$, $F[2.2, 17.4] = 11.39$, $p = .001$, $\eta^2_G = .13$ [.05, .20]) as well as *vice versa* ($BF_{incl} = 3.28$, $F[2.8, 22.4] = 3.97$, $p = .023$, $\eta^2_G = .20$ [.02, .33]).

To investigate the mTEA × CCC interaction when pooled across percepts, we conducted simple-effects mixed ANOVAs separately per CCC. We did not exclude the between-subjects factor Percept from the ANOVAs because the three-way interaction in the main mixed ANOVA was inconclusive (rather than indicating absence of evidence, see above). The results (i.e., the main effects of mTEA in the simple-effects ANOVAs) are visualized in Figure 7. The ANOVA results are summarized in Table S3. For all three CCCs with zero or small tonotopic separation (Z, S_A, and S_B), we found evidence for a medium-to-large mTEA effect. For the CCCs with a large tonotopic separation (L_A and L_B), we found evidence for the absence of an mTEA effect.

In summary, we found evidence for the absence of an effect of the percept on the mTEA effect. Across percepts, we found an effect of mTEA for zero or small tonotopic separations, but not for large tonotopic separations.

General Discussion

The primary focus of our study was to explore the effect of mTEA, the mTEA in dual-electrode CI stimulation on ITD-based left/right discrimination sensitivity. Several tonotopic separations were tested, specified as C1–C2 combinations (CCCs, cf. Figure 2) and selected to inform about the role of channel interactions. Additionally, we explored the origin of the mTEA effect further by comparing the ITD data with data collected for monaural rate-pitch discrimination with a matched experimental setup.

Due to the general heterogeneity of the CI population, our study was designed as a small-*N* within-subjects study with the experimental statistical power focused on the main point of interest, that is, on the individual listener. Note that effects observed in small-*N* studies generalize well as long as the tested hypotheses in principle apply to the entire population and the quality of the collected data is high (Smith & Little, 2018). Despite considerable differences in absolute performance across the five CI listeners tested in

the present study, the effects of the main parameters of interest were quite consistent across individuals.

Effect of the Monaural Temporal Electrode Asynchrony

We hypothesized that mTEA affects dual-electrode ITD discrimination sensitivity, provided that the two stimulating electrodes are closely spaced and, hence, interact already peripherally in the cochlea by stimulating at least partially overlapping auditory-nerve fiber populations. To this end, we tested two CCCs (S_A and S_B) with an average tonotopic separation of about one electrode (i.e., 2–3 mm with the MED-EL electrode array). These CCCs covered the middle part of the electrode array around the C1 electrode. To ensure that peripheral channel interactions indeed occurred with these CCCs, the selection of C2 electrodes was based on forward-masked spatial tuning curves, revealing across-electrode masking of the C1 electrode. Furthermore, to consider the large electrode spacing of MED-EL CIs, we included a control CCC with a tonotopic separation of zero (called Z) in which both C1 and C2 pulse train were presented at the C1 electrode.

In conditions Z and S_B, we indeed found evidence for a large mTEA effect represented as a v- or u-shape with drops in ITD-sensitivity *d'* of up to 1. The mTEA effect was symmetric around 50%, that is, around half the pulse period, indicating that, first, the temporal order of C1 and C2 is not relevant and, second, that channel interactions resulted in an effective increase of the total pulse rate. In both conditions, the mTEA effect was similar in size, suggesting that an electrode separation of about 3 mm is not sufficient to notably reduce the mTEA effect.

In condition S_A, we were not able to obtain any conclusive evidence for either the absence or presence of a delay effect.

We argued that the shape of the mTEA effect is an indicator of the presence of channel interactions. The u-shaped patterns shown in Figures 3 and 6 are overall consistent with the hypothesized effect of channel interactions. This can be best assessed with the zero-separation CCC for which differences between electrodes and tonotopic separation *per se* play no role. For short and long mTEA of 4% and 96%, respectively, for which C1 and C2 pulses are close in time and neurons have maximal time to recover from the C1–C2 pulse pair from the preceding period, ITD sensitivity was only marginally reduced compared to the single-electrode condition. This outcome suggests that with loudness-adjusted dual-electrode stimuli with short or long mTEAs, C1 and C2 pulses almost fully sum up and generate a spiking pattern in the auditory nerve that closely resembles that of the single-electrode condition (cf. Buechel et al., 2018). For medium mTEAs, C1 and C2 pulses are evenly distributed and, thus, the effective pulse rate increases. Hence, the mTEA effect for this condition can also be interpreted from the perspective of rate coding. The largest rate increase can be expected for the 48/52% mTEAs and condition Z, in which, effectively, the pulse rate doubles to 200 pps. For comparison, for an ITD of 500 µs (i.e., matching our average ITD), Ihlefeld et al. (2015) reported a drop in ITD sensitivity by a *d'* of about 1 when the rate was increased from 100 to 200 pps. This corresponds well to the drop in *d'* we found for the 48/52% mTEAs. Furthermore, Egger et al. (2016) reported increased ITD threshold for single-electrode 200-pps low-rate stimulation compared to both small- and especially large-separation 100-pps stimulation. Together, these results indicate that our mTEA effect is a manifestation of the rate limitation found in other studies.

Effect of Tonotopic Separation

Without any tonotopic separation between C1 and C2, an increase in mTEA can be interpreted as an increase in effective pulse rate, as discussed above. In conditions with a clear tonotopic separation, we assumed a similar mechanism, in which the mTEA effect is modulated by the amount of channel interactions. Specifically, we assumed that the amount of interactions decreases with increasing tonotopic separation. Further assuming that across-electrode forward masking is a measure of the amount of channel interactions, we measured forward-masked spatial tuning curves (Nelson et al., 2008, 2011) to derive CCCs eliciting minimal forward masking (i.e., L_A and L_B). We tested these CCCs for an mTEA effect. Indeed, for L_A, we found evidence for the absence of an mTEA effect. ITD sensitivity was nearly as good as that predicted from optimal integration across electrodes (cf. $\Delta d'$ scores around 0 in Figure 6). For L_B, however, we found only inconclusive evidence. For mTEAs of 4% and 32%, the sensitivity was similar to that in L_A (cf. Figure 6). For all other mTEAs, the sensitivity dropped without any v/u-shaped pattern. There is, thus, no evidence for the general absence of mTEA in this condition. However, given the lack of a v/u-shaped pattern, the origin of this mTEA effect likely differs from the effect we found for small and zero-separations CCCs (cf. section "Effect of the monaural temporal electrode asynchrony"). One possibility is spurious masking of the basal C2 by C1, which was not captured by the fmSTCs.

Our results suggest that electrodes separated by about 10 mm in the apical half and by 7–8 mm in the basal half of the CI electrode array stimulate largely non-overlapping electrode populations. Our results substantially extend the findings of Egger et al. (2016) reporting no change in ITD thresholds when comparing single-electrode versus dual-electrode stimulation with more than 6 mm tonotopic separation and using an mTEA of 50%.

Interaural Time Difference Versus Rate Pitch: The Origin of the mTEA Effect

The effects of mTEA manifest independently at the two ears (e.g., Ihlefeld et al., 2015). In the case of asymmetry between

the ears, the mTEA effect will be dominated by the worse ear, creating a unilateral bottleneck (cf. Anderson et al., 2019, 2022). Hence, to achieve salient binaural cues, two decent monaural electrode–neuron interfaces are necessary. To isolate the mechanism underlying the mTEA effects, we searched for mTEA effects that are specific to binaural hearing and ITD perception. Thus, we compared ITD-based sensitivity with monaural rate-pitch sensitivity. For a fair comparison, the experimental setups were matched (Lindenbeck et al., 2024) and we analyzed only the *relative* dual-electrode sensitivity in terms of $\Delta d'$, that is, contrasting measured dual-electrode sensitivity with the prediction of optimal information integration.

We did not find any clear differences in the mTEA effects for ITD and pitch. Combining the data across the two percepts, we found evidence for an mTEA effect in case of small tonotopic separations (i.e., < 3 mm). These findings appear to contradict the considerable independence of electrodes reported for monaural rate-pitch ranking both with a tonotopic separation of 0.75 mm (Macherey & Carlyon, 2010) and 2.4 mm (Griessner et al., 2021). Furthermore, Griessner et al. (2021) found independence of electrodes via pitch matching of acoustic pure tones to the same electric stimuli in single-sided-deaf CI listeners (tonotopic separation 2.1 mm on average). The most notable difference between the previous studies and the present one is the task (ranking/matching vs. discrimination). In particular, due to our selection of individualized ITDs and FDs in the pretests, our discrimination task was calibrated to be performed around threshold. In contrast to that, the tasks used in Macherey and Carlyon (2010) and Griessner et al. (2021) were performed more suprathreshold. Furthermore, it is unclear to what extent the CI listeners in the different studies focused on one electrode or integrated the information across electrodes to solve the tasks, regardless of the actual amount of peripheral independence.

For wide separations (i.e., > 7 mm), we also found evidence for the absence of difference in the mTEA effect between pitch and ITD percepts. This pattern is consistent with the idea that for both percepts channel interactions are responsible for the observed interplay between the effects of tonotopic separation and mTEA.

Taken together, the absence of a consistent distinction in binaural stimulation compared to monaural stimulation suggests a monaural origin for the mTEA effect that manifests independently, however, potentially to a different degree, in the two ears.

Summary and Conclusions

We studied dual-electrode ITD-based left/right discrimination sensitivity in bilaterally implanted CI listeners as a function of two main parameters: the mTEA and the tonotopic separation of stimulation electrodes. The basic stimulus was unmodulated 100-pps pulse trains. A centrally placed interaural electrode pair (C1) was combined with either of four surrounding interaural electrode pairs (C2), covering major parts of the electrode array and invoking various degrees of tonotopic separation. The single-electrode interaural pairs resembling the dual-electrode conditions showed homogeneous ITD sensitivity (d').

The results confirmed our overall hypothesis of an interaction between the effects of mTEA and tonotopic separation. For tonotopic separations smaller than 3 mm, we found clear evidence for an mTEA effect in two out of three conditions (i.e., T and S_B, but not S_A). For tonotopic separations of more than 7 mm (i.e., conditions L_A and L_B), we found little (condition L_B) to no (condition L_A) evidence for an mTEA effect; ITD sensitivity was nearly as good as predicted from optimal integration across electrodes.

To shed light on the origin of the mTEA effect, we compared ITD discrimination sensitivity with *monaural* rate-pitch discrimination sensitivity and did not find systematic differences between these two percepts. In fact, when pooled across all electrode-pair combinations, we even found evidence for the *absence* of the percept effect.

When pooled across percepts, we found evidence for an mTEA effect for all tonotopic separations smaller than 3 mm and evidence for the absence of an mTEA effect for tonotopic separations larger than about 7 mm. We conclude that channel interactions limit ITD-based perception for tonotopic separations up to at least about 3 mm. Future studies may attempt to determine the exact tonotopic separation (between 3 and 7 mm) for which the mTEA effect vanishes.

Overall, both the decreasing size of the mTEA effect with increasing tonotopic separation and the mTEA effect similarity between the binaural and monaural stimulation suggest that the mTEA effect is a *peripheral monaural* effect even when tested in a binaural ITD task. As practical implication, our results suggest that the stimulation delay between nearby and successively to encode a sound event-stimulated electrodes should be as short as possible.

Acknowledgments

We thank our CI listeners for their patience and commitment while conducting the experiments. We further thank Michael Mihocic, Tobias Harvey, and Marie-Luise Augsten for assisting with data collection as well as Michael Smithson for extensive discussions on confidence intervals for effect sizes. We thank the Institute of Ion Physics and Applied Physics of the Leopold-Franzens-University, Innsbruck, Austria, for providing the equipment for direct electric stimulation.

Author Contributions

MJL and BL primarily conceptualized and designed the study. MJL collected the data and performed the statistical analysis. MJL wrote the first draft of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

Data Availability Statement

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declaration of Conflicting Interests

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: MJL was supported by a DOC Fellowship of the Austrian Academy of Sciences (A-25606) and by the Dr. Anton Oelzelt-Newin Fund of the Austrian Academy of Sciences (A-100315). Additional support was provided to all authors by the Austrian Agency for International Cooperation in Education and Research (MULT 07/2020). Open access funding was provided by the Open Access Fund of the Austrian Academy of Sciences.

ORCID iDs

Martin J. Lindenbeck https://orcid.org/0000-0001-7779-4767
Bernhard Laback https://orcid.org/0000-0003-0929-6787

Supplemental material

Supplemental material for this article is available online.

Notes

1. Portions of this work were presented at the Conference on Implantable Auditory Prostheses, July 2021, Lake Tahoe, CA, the 48th Annual Meeting of the German Acoustical Society, March 2022, Stuttgart, Germany, and the 19th International Symposium on Hearing, June 2022, Lyon, France.
2. The exact delays deviated negligibly due to technical reasons. The exact delays were 3.96%, 31.87%, 47.82%, 52.18%, 68.13%, and 96.04%.
3. As masker and target had the same pulse rate, the target might have been perceived as a continuation of the masker in case both pulse trains were presented to the same electrode. This might have increased the sharpness of the tuning curves (Cosentino et al., 2015; Moore et al., 1984; Neff, 1985). Yet, the selection of C2 electrodes was not based on sharpness.
4. CI24 was the only listener with an electrode spacing of two electrodes (instead of a spacing of one, i.e., electrode eight instead of seven) in the S_B condition and for the right ear (cf. Table 2). The electrode spacing was increased due to absent masking of the closer electrode seven in the tuning curve.
5. For the zero-separation condition, this score was also the C1-electrode-only *d'*, hence, in this condition $\overline{d'} = d'_{\text{single, z}}$.
6. ITD-discrimination sensitivity has been shown to depend on loudness (e.g., Egger et al., 2016, 2017).
7. Note that this prediction is based on the probably simplistic assumption that *d'* rises linearly with level, at least around the comfortable level. Yet, any departure from this assumption should at least be similar for the two percepts compared (ITD and rate pitch) and, thus, not affect their relative comparison.

References

Amrhein, V., Greenland, S., & McShane, B. (2019). Scientists rise up against statistical significance. *Nature*, *567*(7748), 305–307. https://doi.org/10.1038/d41586-019-00857-9

Anderson, S. R., Kan, A., & Litovsky, R. Y. (2019). Asymmetric temporal envelope encoding: Implications for within- and across-ear envelope comparison. *The Journal of the Acoustical Society of America*, *146*(2), 1189–1206. https://doi.org/10.1121/1.5121423

Anderson, S. R., Kan, A., & Litovsky, R. Y. (2022). Asymmetric temporal envelope sensitivity: Within- and across-ear envelope comparisons in listeners with bilateral cochlear implants. *The Journal of the Acoustical Society of America*, *152*(6), 3294–3312. https://doi.org/10.1121/10.0016365

Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior Research Methods*, *37*(3), 379–384. https://doi.org/10.3758/BF03192707

Bendor, D., & Wang, X. (2005). The neuronal representation of pitch in primate auditory cortex. *Nature*, *436*(7054), 1161–1165. https://doi.org/10.1038/nature03867

Bernstein, J. G. W., Jensen, K. K., Stakhovskaya, O. A., Noble, J. H., Hoa, M., Kim, H. J., & Goupell, M. J. (2021). Interaural place-of-stimulation mismatch estimates using CT scans and binaural perception, but not pitch, are consistent in cochlear-implant users. *The Journal of Neuroscience*, *41*(49), 10161–10178. https://doi.org/10.1523/JNEUROSCI.0359-21.2021

Blauert, J. (1997). *Spatial hearing: The psychophysics of human sound localization* (Rev. ed). MIT Press.

Boulet, J., White, M., & Bruce, I. C. (2016). Temporal considerations for stimulating spiral ganglion neurons with cochlear implants. *Journal of the Association for Research in Otolaryngology*, *17*(1), 1–17. https://doi.org/10.1007/s10162-015-0545-5

Brughera, A., Dunai, L., & Hartmann, W. M. (2013). Human interaural time difference thresholds for sine tones: The high-frequency limit. *The Journal of the Acoustical Society of America*, *133*(5), 2839–2855. https://doi.org/10.1121/1.4795778

Buechel, B. D., Hancock, K. E., Chung, Y., & Delgutte, B. (2018). Improved neural coding of ITD with bilateral cochlear implants by introducing short inter-pulse intervals. *Journal of the Association for Research in Otolaryngology*, *19*(6), 681–702. https://doi.org/10.1007/s10162-018-00693-0

Buell, T. N., & Hafter, E. R. (1991). Combination of binaural information across frequency bands. *The Journal of the Acoustical Society of America*, *90*(4), 1894–1900. https://doi.org/10.1121/1.401668

Buell, T. N., & Trahiotis, C. (1993). Interaural temporal discrimination using two sinusoidally amplitude-modulated, high-frequency tones: Conditions of summation and interference. *The Journal of the Acoustical Society of America*, *93*(1), 480–487. https://doi.org/10.1121/1.405628

Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist*, *49*(12), 997–1003. https://doi.org/10.1037/0003-066X.49.12.997

Cosentino, S., Deeks, J. M., & Carlyon, R. P. (2015). Procedural factors that affect psychophysical measures of spatial selectivity in cochlear implant users. *Trends in Hearing*, *19*, 233121651560706. https://doi.org/10.1177/2331216515607067

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology*, *1*(1), 42–45. https://doi.org/10.20982/tqmp.01.1.p042

Cumming, G. (2014). The new statistics: Why and how. *Psychological Science*, *25*(1), 7–29. https://doi.org/10.1177/0956797613504966

de Balthasar, C., Boëx, C., Cosendai, G., Valentini, G., Sigrist, A., & Pelizzone, M. (2003). Channel interactions with high-rate biphasic electrical stimulation in cochlear implant subjects. *Hearing Research*, *182*(1–2), 77–87. https://doi.org/10.1016/S0378-5955(03)00174-6

Dienes, Z., & Mclatchie, N. (2018). Four reasons to prefer Bayesian analyses over significance testing. *Psychonomic Bulletin & Review*, *25*(1), 207–218. https://doi.org/10.3758/s13423-017-1266-z

Eddington, D. K., Dobelle, W. H., Brackmann, D. E., Mladejovsky, M. G., & Parkin, J. L. (1978). Auditory prostheses research with multiple channel intracochlear stimulation in man. *The Annals of Otology, Rhinology, and Laryngology*, *87*(6 Pt 2), 1–39.

Egger, K., Majdak, P., & Laback, B. (2016). Channel interaction and current level affect across-electrode integration of interaural time differences in bilateral cochlear-implant listeners. *Journal of the Association for Research in Otolaryngology*, *17*(1), 55–67. https://doi.org/10.1007/s10162-015-0542-8

Egger, K., Majdak, P., & Laback, B. (2017). Binaural timing information in electric hearing at low rates: Effects of inaccurate encoding and loudness. *The Journal of the Acoustical Society of America*, *141*(5), 3164–3174. https://doi.org/10.1121/1.4982888

Favre, E., & Pelizzone, M. (1993). Channel interactions in patients using the Ineraid multichannel cochlear implant. *Hearing Research*, *66*(2), 150–156. https://doi.org/10.1016/0378-5955(93)90136-O

Francart, T., Lenssen, A., Büchner, A., Lenarz, T., & Wouters, J. (2015). Effect of channel envelope synchrony on interaural time difference sensitivity in bilateral cochlear implant listeners. *Ear and Hearing*, *36*(4), e199–e206. https://doi.org/10.1097/AUD.0000000000000152

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. John Wiley.

Griessner, A., Schatzer, R., Steixner, V., Rajan, G. P., Zierhofer, C., & Távora-Vieira, D. (2021). Temporal pitch perception in cochlear-implant users: Channel independence in apical cochlear regions. *Trends in Hearing*, *25*, 233121652110206. https://doi.org/10.1177/23312165211020645

Grothe, B., Pecka, M., & McAlpine, D. (2010). Mechanisms of sound localization in mammals. *Physiological Reviews*, *90*(3), 983–1012. https://doi.org/10.1152/physrev.00026.2009

Hartmann, R., Topp, G., & Klinke, R. (1984). Discharge patterns of cat primary auditory fibers with electrical stimulation of the cochlea. *Hearing Research*, *13*(1), 47–62. https://doi.org/10.1016/0378-5955(84)90094-7

Hu, H., & Dietz, M. (2015). Comparison of interaural electrode pairing methods for bilateral cochlear implants. *Trends in Hearing*, *19*, 1–22. https://doi.org/10.1177/2331216515617143

Huynh, H., & Feldt, L. S. (1970). Conditions under which mean square ratios in repeated measurements designs have exact F-distributions. *Journal of the American Statistical Association*, *65*(332), 1582–1589. https://doi.org/10.1080/01621459.1970.10481187

Ihlefeld, A., Carlyon, R. P., Kan, A., Churchill, T. H., & Litovsky, R. Y. (2015). Limitations on monaural and binaural temporal processing in bilateral cochlear implant listeners. *Journal of the Association for Research in Otolaryngology*, *16*(5), 641–652. https://doi.org/10.1007/s10162-015-0527-7

Ihlefeld, A., Kan, A., & Litovsky, R. Y. (2014). Across-frequency combination of interaural time difference in bilateral cochlear implant listeners. *Frontiers in Systems Neuroscience*, *8*, https://doi.org/10.3389/fnsys.2014.00022

JASP Team. (2023). *JASP* (Version 0.17) [Computer software]. https://jasp-stats.org/.

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*(430), 773–795. https://doi.org/10.1080/01621459.1995.10476572

Keysers, C., Gazzola, V., & Wagenmakers, E.-J. (2020). Using Bayes factor hypothesis testing in neuroscience to establish evidence of absence. *Nature Neuroscience*, *23*(7), 788–799. https://doi.org/10.1038/s41593-020-0660-4

Klein-Hennig, M., Dietz, M., Hohmann, V., & Ewert, S. D. (2011). The influence of different segments of the ongoing envelope on sensitivity to interaural time delays. *The Journal of the Acoustical Society of America*, *129*(6), 3856–3872. https://doi.org/10.1121/1.3585847

Klein, S. A. (2001). Measuring, estimating, and understanding the psychometric function: A commentary. *Perception & Psychophysics*, *63*(8), 1421–1455. https://doi.org/10.3758/BF03194552

Klingel, M., & Laback, B. (2022). Reweighting of binaural localization cues in bilateral cochlear-implant listeners. *Journal of the Association for Research in Otolaryngology : JARO*, *23*(1), 119–136. https://doi.org/10.1007/s10162-021-00821-3

Kong, Y.-Y., Deeks, J. M., Axon, P. R., & Carlyon, R. P. (2009). Limits of temporal pitch in cochlear implants. *The Journal of the Acoustical Society of America*, *125*(3), 1649–1657. https://doi.org/10.1121/1.3068457

Kreft, H. A., Oxenham, A. J., & Nelson, D. A. (2010). Modulation rate discrimination using half-wave rectified and sinusoidally amplitude modulated stimuli in cochlear-implant users. *The Journal of the Acoustical Society of America*, *127*(2), 656–659. https://doi.org/10.1121/1.3282947

Laback, B., Egger, K., & Majdak, P. (2015). Perception and coding of interaural time differences with bilateral cochlear implants. *Hearing Research*, *322*, 138–150. https://doi.org/10.1016/j.heares.2014.10.004

Laback, B., Majdak, P., & Baumgartner, W.-D. (2007). Lateralization discrimination of interaural time delays in four-pulse sequences in electric and acoustic hearing. *The Journal of the Acoustical Society of America*, *121*(4), 2182–2191. https://doi.org/10.1121/1.2642280

Laback, B., Zimmermann, I., Majdak, P., Baumgartner, W.-D., & Pok, S.-M. (2011). Effects of envelope shape on interaural envelope delay sensitivity in acoustic and electric hearing. *The Journal of the Acoustical Society of America*, *130*(3), 1515–1529. https://doi.org/10.1121/1.3613704

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, *4*, https://doi.org/10.3389/fpsyg.2013.00863

Leek, M. R. (2001). Adaptive procedures in psychophysical research. *Perception & Psychophysics*, *63*(8), 1279–1292. https://doi.org/10.3758/BF03194543

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *The Journal of the Acoustical Society of America*, *49*(2), 467–477. https://doi.org/10.1121/1.1912375

Lindenbeck, M. J., Laback, B., & Majdak, P. (2021, July 12-16). *Sensitivity to interaural time difference and temporal pitch in dual-electrode stimulation: Effects of between-electrode delay, tonotopic separation, and stimulus type* [Poster presentation]. Conference on Implantable Auditory Prostheses, Lake Tahoe, CA, USA.

Lindenbeck, M. J., Laback, B., Majdak, P., & Srinivasan, S. (2020). Temporal-pitch sensitivity in electric hearing with amplitude modulation and inserted pulses with short inter-pulse intervals. *The Journal of the Acoustical Society of America*, *147*(2), 777–793. https://doi.org/10.1121/10.0000610

Lindenbeck, M. J., Majdak, P., Srinivasan, S., & Laback, B. (2023). Pitch discrimination in electric hearing with inconsistent and consistent amplitude-modulation and inter-pulse rate cues. *The Journal of the Acoustical Society of America*, *153*(6), 3268–3280. https://doi.org/10.1121/10.0019452

Long, C. J., Eddington, D. K., Colburn, H. S., & Rabinowitz, W. M. (2003). Binaural sensitivity as a function of interaural electrode position with a bilateral cochlear implant user. *The Journal of the Acoustical Society of America*, *114*(3), 1565–1574. https://doi.org/10.1121/1.1603765

Macherey, O., & Carlyon, R. P. (2010). Temporal pitch percepts elicited by dual-channel stimulation of a cochlear implant. *The Journal of the Acoustical Society of America*, *127*(1), 339–349. https://doi.org/10.1121/1.3269042

Macpherson, E. A., & Middlebrooks, J. C. (2002). Listener weighting of cues for lateral angle: The duplex theory of sound localization revisited. *The Journal of the Acoustical Society of America*, *111*(5), 2219. https://doi.org/10.1121/1.1471898

Majdak, P., Laback, B., & Baumgartner, W.-D. (2006). Effects of interaural time differences in fine structure and envelope on lateral discrimination in electric hearing. *The Journal of the Acoustical Society of America*, *120*(4), 2190–2201. https://doi.org/10.1121/1.2258390

McKay, C. M., & McDermott, H. J. (1996). The perception of temporal patterns for electrical stimulation presented at one or two intracochlear sites. *The Journal of the Acoustical Society of America*, *100*(2), 1081–1092. https://doi.org/10.1121/1.416294

Middlebrooks, J. C., & Green, D. M. (1991). Sound localization by human listeners. *Annual Review of Psychology*, *42*(1), 135–159. https://doi.org/10.1146/annurev.ps.42.020191.001031

Moore, B. C. J., Glasberg, B. R., & Roberts, B. (1984). Refining the measurement of psychophysical tuning curves. *The Journal of the Acoustical Society of America*, *76*(4), 1057–1066. https://doi.org/10.1121/1.391425

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorials in Quantitative Methods for Psychology*, *4*(2), 61–64. https://doi.org/10.20982/tqmp.04.2.p061

Neff, D. L. (1985). Stimulus parameters governing confusion effects in forward masking. *The Journal of the Acoustical Society of America*, *78*(6), 1966–1976. https://doi.org/10.1121/1.392653

Nelson, D. A., Donaldson, G. S., & Kreft, H. (2008). Forward-masked spatial tuning curves in cochlear implant users. *The Journal of the Acoustical Society of America*, *123*(3), 1522–1543. https://doi.org/10.1121/1.2836786

Nelson, D. A., Kreft, H. A., Anderson, E. S., & Donaldson, G. S. (2011). Spatial tuning curves from apical, middle, and basal electrodes in cochlear implant users. *The Journal of the Acoustical Society of America*, *129*(6), 3916–3933. https://doi.org/10.1121/1.3583503

Noel, V. A., & Eddington, D. K. (2013). Sensitivity of bilateral cochlear implant users to fine-structure and envelope interaural time differences. *The Journal of the Acoustical Society of America*, *133*(4), 2314–2328. https://doi.org/10.1121/1.4794372

Oberfeld, D., & Franke, T. (2013). Evaluating the robustness of repeated measures analyses: The case of small sample sizes and nonnormal data. *Behavior Research Methods*, *45*(3), 792–812. https://doi.org/10.3758/s13428-012-0281-2

Olejnik, S., & Algina, J. (2003). Generalized eta and omega squared statistics: Measures of effect size for some common research designs. *Psychological Methods*, *8*(4), 434–447. https://doi.org/10.1037/1082-989X.8.4.434

Poon, B. B., Eddington, D. K., Noel, V., & Colburn, H. S. (2009). Sensitivity to interaural time difference with bilateral cochlear implants: Development over time and effect of interaural electrode spacing. *The Journal of the Acoustical Society of America*, *126*(2), 806–815. https://doi.org/10.1121/1.3158821

Rouder, J. N., Engelhardt, C. R., McCabe, S., & Morey, R. D. (2016). Model comparison in ANOVA. *Psychonomic Bulletin & Review*, *23*(6), 1779–1786. https://doi.org/10.3758/s13423-016-1026-5

Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes factors for ANOVA designs. *Journal of Mathematical Psychology*, *56*(5), 356–374. https://doi.org/10.1016/j.jmp.2012.08.001

Rouder, J. N., Morey, R. D., Verhagen, J., Swagman, A. R., & Wagenmakers, E.-J. (2017). Bayesian analysis of factorial designs. *Psychological Methods*, *22*(2), 304–321. https://doi.org/10.1037/met0000057

Smith, Z. M., Delgutte, B., & Oxenham, A. J. (2002). Chimaeric sounds reveal dichotomies in auditory perception. *Nature*, *416*(6876), 87–90. https://doi.org/10.1038/416087a

Smith, P. L., & Little, D. R. (2018). Small is beautiful: In defense of the small-N design. *Psychonomic Bulletin & Review*, *25*(6), 2083–2101. https://doi.org/10.3758/s13423-018-1451-8

Srinivasan, S., Laback, B., Majdak, P., & Arnoldner, C. (2020). Improving interaural time difference sensitivity using short interpulse intervals with amplitude-modulated pulse trains in bilateral cochlear implants. *Journal of the Association for Research in Otolaryngology*, *21*(1), 105–120. https://doi.org/10.1007/s10162-020-00743-6

Thakkar, T., Kan, A., Jones, H. G., & Litovsky, R. Y. (2018). Mixed stimulation rates to improve sensitivity of interaural timing differences in bilateral cochlear implant listeners. *The Journal of the Acoustical Society of America*, *143*(3), 1428–1440. https://doi.org/10.1121/1.5026618

Thakkar, T., Kan, A., & Litovsky, R. Y. (2023). Lateralization of interaural time differences with mixed rates of stimulation in bilateral cochlear implant listeners. *The Journal of the Acoustical Society of America*, *153*(3), 1912–1923. https://doi.org/10.1121/10.0017603

van den Bergh, D., van Doorn, J., Marsman, M., Draws, T., Van Kesteren, E.-J., Derks, K., & Wagenmakers, E.-J. (2020). A tutorial on conducting and interpreting a Bayesian ANOVA in JASP. *L'Année Psychologique*, *120*(1), 73–96. https://doi.org/10.3917/anpsy1.201.0073

van den Bergh, D., Wagenmakers, E.-J., & Aust, F. (2023). Bayesian repeated-measures ANOVA: An updated methodology

implemented in JASP. *Advances in Methods and Practices in Psychological Science, 6*(2), 25152459231168024. https://doi.org/10.1177/25152459231168024

Van Den Honert, C., & Stypulkowski, P. H. (1987). Temporal response patterns of single auditory nerve fibers elicited by periodic electrical stimuli. *Hearing Research*, *29*(2–3), 207–222. https://doi.org/10.1016/0378-5955(87)90168-7

Van Eeckhoutte, M., Spirrov, D., & Francart, T. (2018). Comparison between adaptive and adjustment procedures for binaural loudness balancing. *The Journal of the Acoustical Society of America*, *143*(6), 3720–3729. https://doi.org/10.1121/1.5042522

van Hoesel, R. J. M. (2007). Sensitivity to binaural timing in bilateral cochlear implant users. *The Journal of the Acoustical Society of America*, *121*(4), 2192–2206. https://doi.org/10.1121/1.2537300

van Hoesel, R. J. M., Jones, G. L., & Litovsky, R. Y. (2009). Interaural time-delay sensitivity in bilateral cochlear implant users: Effects of pulse rate, modulation rate, and place of stimulation. *Journal of the Association for Research in Otolaryngology*, *10*(4), 557–567. https://doi.org/10.1007/s10162-009-0175-x

Wagenmakers, E.-J., Verhagen, J., Ly, A., Matzke, D., Steingroever, H., Rouder, J. N., & Morey, R. D. (2017). The need for Bayesian hypothesis testing in psychological science. In S. O. Lilienfeld & I. D. Waldman (Eds.), *Psychological science under scrutiny* (pp. 123–138). John Wiley & Sons, Inc. https://doi.org/10.1002/9781119095910.ch8

Wetzels, R., Grasman, R. P. P. P., & Wagenmakers, E.-J. (2012). A default Bayesian hypothesis test for ANOVA designs. *The American Statistician*, *66*(2), 104–111. https://doi.org/10.1080/00031305.2012.695956

Wightman, F. L., & Kistler, D. J. (1992). The dominant role of low-frequency interaural time differences in sound localization. *The Journal of the Acoustical Society of America*, *91*(3), 1648–1661. https://doi.org/10.1121/1.402445

Yoon, Y., Li, Y., Kang, H.-Y., & Fu, Q.-J. (2011). The relationship between binaural benefit and difference in unilateral speech recognition performance for bilateral cochlear implant users. *International Journal of Audiology*, *50*(8), 554–565. https://doi.org/10.3109/14992027.2011.580785

Zeng, F.-G. (2022). Celebrating the one millionth cochlear implant. *JASA Express Letters*, *2*(7), 077201. https://doi.org/10.1121/10.0012825

Zeng, F.-G., Nie, K., Liu, S., Stickney, G., Del Rio, E., Kong, Y.-Y., & Chen, H. (2004). On the dichotomy in auditory perception between temporal envelope and fine structure cues (L). *The Journal of the Acoustical Society of America*, *116*(3), 1351–1354. https://doi.org/10.1121/1.1777938

Supplementary Material

Manuscript title:	Effects of Monaural Temporal Electrode Asynchrony and Channel Interactions in Bilateral and Unilateral Cochlear-Implant Stimulation
Authors:	Martin J. Lindenbeck[1], Piotr Majdak, and Bernhard Laback
Date:	September 2, 2024

[1] martin.lindenbeck@oeaw.ac.at

Statistical analysis

Bayesian hypothesis testing Although the Bayes Factor (BF) is a continuous measure of evidence (in contrast to *p* values; Lakens, 2022), we classified BFs in ordinal categories for ease of interpretation (broadly following van Doorn et al., 2019). Normally, BFs between 3 and 10 are considered as moderate evidence for H1 over H0 (i.e., "evidence of presence") and BFs greater than 10 as strong evidence. Furthermore, the inverse thresholds are used to establish the evidence of H0 over H1 ("evidence of absence"; moderate: 0.33 to 0.10; strong: < 0.10). BFs between 0.33 and 3 are considered anecdotal (i.e., "absence of evidence"). To reduce the statistical complexity, we introduced only three categories of evidence: evidence of presence (BF >3), evidence of absence (BF < 0.33), and inconclusive evidence (BF between 0.33 and 3). Note that by introducing ordinal categories for the BF, both Type I and Type II errors could have also occurred with BFs (Lakens, 2023) which is another reason why we still report NHST statistics and *p* values.

The Bayesian approach to rmANOVA relies on comparison of the predictive performance of ANOVA models. For a certain rmANOVA, the null model and all possible main-effect(s)-only and interaction models are assigned equal prior model probabilities. The null model includes both random intercepts and random slopes (van den Bergh, Wagenmakers, & Aust, 2022). Subsequently, given the data, the predictive performance is assessed via posterior model probabilities. Note that this Bayesian approach automatically favors parsimony over complexity (van den Bergh et al., 2020) and, hence, protects against overfitting. Subsequently, evidence for main effects and interactions is quantified by comparing the posterior probabilities of models that differ by only one main effect or interaction, respectively (matched-model analysis of effects; cf. Keysers, Gazzola, & Wagenmakers, 2020; van den Bergh et al., 2020).

For all Bayesian rmANOVAs, the normality-of-residuals assumption was confirmed via visual inspection of Q-Q plots. Note that in the Bayesian framework 95% credible intervals can be constructed for each residual separately. Normality was assumed when all credible intervals intersected the line signaling normality in the Q-Q plot.

Bayesian effect-size estimates similar to null-hypothesis significance-testing η^2 estimates are still in development (Keysers et al., 2020).

Null-hypothesis significance testing and confidence intervals Somewhat similar to Bayesian hypothesis testing, there are also NHST procedures to collect evidence for the H0 (i.e., equivalence tests; e.g., Lakens, 2017). However, to date, these tests are not available for comparisons of more than two conditions (i.e., ANOVAs).

The η_G^2 confidence intervals were derived from non-central *F* distributions using the *F* statistics, Huynh-Feldt-corrected *p* values and degrees of freedom (for more details, see Smithson, 2003). To estimate the non-centrality parameter, the *F*-statistic and the η_G^2-associated degrees of freedom (*df*) are needed. However, for an rmANOVA, only the partial η^2 (η_p^2) can directly be computed from the rmANOVA output using $\eta_p^2 = \frac{F \cdot df_1}{F \cdot df_1 + df_{2,p}}$, where df_1 and $df_{2,p}$ are the *F*-test's nominator and denominator degrees of freedom, respectively. Crucially, η_G^2 differs from η_p^2 in that it contains a larger portion of (unexplained, non-manipulated) variance in the denominator, hence, $\eta_G^2 \geq \eta_p^2$ (see Bakeman, 2005; Olejnik & Algina, 2003). Consequently, with additional variance from other rmANOVA terms included, $df_{2,G} \geq df_{2,p}$. To estimate $df_{2,G}$, we assumed that η_G^2 can be computed similarly to η_p^2, $\eta_G^2 = \frac{F \cdot df_1}{F \cdot df_1 + df_{2,G}}$, and calculated $df_{2,G}$ using the second formula. df_1, $df_{2,G}$, and the *F* statistic were then used to estimate the non-centrality parameter. To this end, for each term of an rmANOVA, the non-centrality parameter was estimated with an iterative procedure using the ncpci() MATLAB function of the Measures of Effect Size Toolbox (Hentschke, 2017/2022; Hentschke & Stüttgen, 2011). Note that, since the η_G^2 confidence intervals were derived from the *p* values, these two measures are inherently linked.

Furthermore, note that a non-zero η_G^2 might be accompanied by a $BF_{incl} < 0.33$ indicating evidence for the absence of an effect (i.e., $\eta_G^2 = 0$). In such cases, the upper bound of the η_G^2 confidence interval can be thought of as a worst-case approximation for a Type II error (Keysers et al., 2020).

Tables

Table S1. Results of the one-way rmANOVAs conducted to assess the mTEA effect separately per CCC. Zero tonotopic separation: CCC Z. Small tonotopic separation: CCCs S_A and S_B. Large tonotopic separation: CCCs L_A and L_B. Per rmANOVA: BF_{incl}…Bayes factor; F…test statistic; df_{mTEA}…Huynh-Feldt-adjusted nominator degrees of freedom; df_{error}…Huynh-Feldt-adjusted denominator degrees of freedom; p…Huynh-Feldt-adjusted significance level; η_G^2…effect size estimate [90% confidence intervals].

CCC	BF_{incl}	F	df_{mTEA}	df_{error}	p	η_G^2 [5%, 95%]
Z	6.40	3.87	4.7	18.6	.016	.36 [.07, .47]
S_A	1.04	2.02	5.0	20.0	.120	.21 [.00, .31]
S_B	191.73	9.26	2.1	8.3	.007	.32 [.11, .46]
L_A	0.18	0.38	4.3	17.2	.835	.03 [.00, .05]
L_B	1.84	2.83	5.0	20.0	.043	.12 [.01, .18]

Table S2. Results of the Percept × mTEA interaction in the two-way mixed ANOVAs conducted to assess the effect of the percept (ITD or rate pitch) on the mTEA effect separately per CCC. All other aspects as in Table S1.

CCC	BF_{incl}	F	df_{mTEA}	df_{error}	p	η_G^2 [5%, 95%]
Z	0.26	0.81	3.7	30.0	.524	.03 [.00, .07]
S_A	3.51	3.07	2.6	21.1	.056	.18 [.00, .32]
S_B	0.18	0.62	3.0	23.8	.756	.02 [.00, .05]
L_A	0.39	1.04	4.0	32.3	.402	.05 [.00, .10]
L_B	1.46	2.23	5.0	40.0	.070	.04 [.00, .06]

Table S3. Results of the mTEA main effect in the two-way mixed ANOVAs conducted to assess the effect of the percept on the mTEA effect separately per CCC. All other aspects as in Table S1.

CCC	BF_{incl}	F	df_{mTEA}	df_{error}	p	η_G^2 [5%, 95%]
Z	2591.64	8.64	3.7	30.0	< .001	.24 [.11, .33]
S_A	93.82	6.84	2.6	21.1	.003	.33 [.10, .46]
S_B	12055.61	9.53	3.0	23.8	< .001	.27 [.12, .37]
L_A	0.10	0.42	4.0	32.3	.798	.02 [.00, .04]
L_B	0.26	1.40	5.0	40.0	.246	.02 [.00, .04]

References

Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Behavior Research Methods*, *37*(3), 379–384. doi: 10.3758/BF03192707

Hentschke, H. (2022). *Measures of Effect Size (MES) Toolbox* [Matlab]. Retrieved from https://github.com/hhentschke/measures-of-effect-size-toolbox (Original work published 2017)

Hentschke, H., & Stüttgen, M. C. (2011). Computation of measures of effect size for neuroscience data sets: Effect size toolbox. *European Journal of Neuroscience*, *34*(12), 1887–1894. doi: 10.1111/j.1460-9568.2011.07902.x

Keysers, C., Gazzola, V., & Wagenmakers, E.-J. (2020). Using Bayes factor hypothesis testing in neuroscience to establish evidence of absence. *Nature Neuroscience*, *23*(7), 788–799. doi: 10.1038/s41593-020-0660-4

Lakens, Daniël. (2017). Equivalence Tests: A Practical Primer for *t* Tests, Correlations, and Meta-Analyses. *Social Psychological and Personality Science*, *8*(4), 355–362. doi: 10.1177/1948550617697177

Lakens, Daniël. (2022). Why P values are not measures of evidence. *Trends in Ecology & Evolution*, *37*(4), 289–290. doi: 10.1016/j.tree.2021.12.006

Lakens, Daniel. (2023, April 13). The 20% Statistician: Preventing common misconceptions about Bayes Factors. Retrieved April 24, 2023, from The 20% Statistician website: http://daniellakens.blogspot.com/2023/04/preventing-common-misconceptions-about.html

Olejnik, S., & Algina, J. (2003). Generalized Eta and Omega Squared Statistics: Measures of Effect Size for Some Common Research Designs. *Psychological Methods*, *8*(4), 434–447. doi: 10.1037/1082-989X.8.4.434

Smithson, M. (2003). *Confidence intervals*. Thousand Oaks, CA: Sage.

van den Bergh, D., van Doorn, J., Marsman, M., Draws, T., Van Kesteren, E.-J., Derks, K., … Wagenmakers, E.-J. (2020). A Tutorial on Conducting and Interpreting a Bayesian ANOVA in JASP. *L'Année Psychologique*, *Vol. 120*(1), 73–96. doi: 10.3917/anpsy1.201.0073

van den Bergh, D., Wagenmakers, E.-J., & Aust, F. (2022). *Bayesian Repeated-Measures ANOVA: An Updated Methodology Implemented in JASP* [Preprint]. PsyArXiv. doi: 10.31234/osf.io/fb8zn

van Doorn, J., van Den Bergh, D., Bohm, U., Dablander, F., Derks, K., Draws, T., … Wagenmakers, E.-J. (2019). *The JASP Guidelines for Conducting and Reporting a Bayesian Analysis* [Preprint]. PsyArXiv. doi: 10.31234/osf.io/yqxfr

Channel Interactions with Pulse Rate, Amplitude Modulation and Short Inter-Pulse Interval Cues in Unilateral Dual-Electrode Stimulation

The following manuscript was submitted on 16 April 2025 to Trends in Hearing:

Lindenbeck, M. J., Majdak, P., & Laback, B. (2025). Effects of Monaural Temporal Electrode Asynchrony on Temporal Pitch Discrimination with Cochlear Implants [manuscript submitted for publication]. *Acoustics Research Institute, Austrian Academy of Sciences*.

Author contributions

MJL and BL primarily conceptualized and designed the study. MJL collected the data and performed the statistical analysis. MJL wrote the first draft of the manuscript. All authors contributed to manuscript revision, read and approved the submitted version.

Effects of Monaural Temporal Electrode Asynchrony on Temporal Pitch Discrimination with Cochlear Implants[1]

Martin J. Lindenbeck[2], Piotr Majdak, and Bernhard Laback[3]

Acoustics Research Institute, Austrian Academy of Sciences

Dominikanerbastei 16, A-1010 Vienna, AUSTRIA

ORCID

Martin J. Lindenbeck: 0000-0001-7779-4767; martin.lindenbeck@oeaw.ac.at

Piotr Majdak: 0000-0003-1511-6164; piotr.majdak@oeaw.ac.at

Bernhard Laback: 0000-0003-0929-6787; bernhard.laback@oeaw.ac.at

March 6, 2025

Keywords

pitch discrimination; channel interaction; electrode separation; short inter-pulse interval; rate limitation; auditory nerve; phenomenological model

Abstract

Cochlear-implant (CI) listeners show impaired pitch perception compared to normal-hearing listeners. One of the factors limiting pitch sensitivity in multi-electrode as compared to single-electrode stimulation can be intracochlear interactions of electrode signals (i.e., channels). To shed more light on this issue, we measured temporal-pitch discrimination sensitivity for loudness-balanced single- and dual-electrode stimuli with various spatio-temporal configurations in listeners with MED-EL CIs. We hypothesized a link between pitch sensitivity and tonotopic electrode separation as well as (monaural) temporal electrode asynchrony (mTEA), the latter potentially resulting in various combinations of inter-pulse intervals (IPIs) in the compound stimuli received by the auditory nerve. Stimulus types included low-rate (LR, 100 pps), high-rate with 100-Hz amplitude modulation (HR, 1000 pps), and SIPI, a special type of HR including extra pulses inserted with short IPIs at the modulation peaks. Single-electrode sensitivity was best for LR stimuli and worst for HR stimuli. With dual-electrode stimulation, there was a detrimental mTEA effect for tonotopic separations below 2.2 mm but not for separations of 7.1 mm and more. This pattern was largely consistent across stimulus types. Simulations with a phenomenological model of an auditory-nerve fiber [Takanen & Seeber, Trends Hear 26, 233121652211170 (2022)] suggest channel interactions to be the underlying cause, manifesting for LR stimulation as changes in effective rate (consistent with IPI-based temporal-pitch theories) and for HR stimulation as coarser modulation coding due to spiking at off-peak carrier pulses. SIPI stimulation combined beneficial properties of LR and HR stimulation, providing prominent modulation coding combined with a moderate channel interaction effect on the effective modulation rate. Taken together, these results may guide strategies to optimally distribute stimulation pulses across electrodes and time with respect to temporal-pitch sensitivity.

Introduction

Pitch is the perceptual correlate of acoustic waveform periodicity and a ubiquitous part of everyday life. It is essential for speech perception, enabling speaker identification (e.g., Fu et al., 2004; Vongphoe & Zeng, 2005), prosody understanding (e.g., Cutler et al., 1997), and speech understanding in noise (e.g., Miller et al., 2010). In music, pitch cues allow for the recognition of melodies and form the basis of harmony (e.g., Houtsma, 1995; Parncutt, 1989). Additionally, pitch cues are vital for auditory scene analysis, particularly the segregation of sound sources (e.g., Darwin, 2005; Darwin & Carlyon, 1995).

In normal hearing (NH), pitch may be conveyed both via the tonotopic place of stimulation (i.e., the place code) and by means of phased-locked periodic auditory-nerve activity at a given place of stimulation (i.e., the temporal code). For a review see, for example, Oxenham (2018). Even in basic tasks such as pure-tone frequency discrimination, NH listeners can discriminate tones differing in frequency by less than 1% (e.g., Dai & Micheyl, 2011).

Yet, pitch perception is rather different in cochlear-implant (CI) listeners. CIs are auditory prostheses originally designed to provide speech understanding in the profoundly deaf or hearing-impaired (see, e.g., Wilson & Dorman, 2008; Zeng, 2022). In frequency or rate discrimination, respectively, CI listeners perform much worse than their NH counterparts with just-noticeable differences around 6%, even when stimulating in a highly-controlled manner with research interfaces instead of clinical processors and at a single intracochlear electrode only (e.g., Moore & Carlyon, 2005). With clinical processors and multi-electrode stimulation, just-noticeable differences can be as high as 70% (Gaudrain & Başkent, 2018). The discrepancy between NH and single-electrode CI sensitivity is thought to be a result of poor place-pitch (i.e., spectral) sensitivity and CI listeners relying largely on temporal pitch cues similar to those derived from unresolved harmonics in NH (Carlyon et al., 2002;

Carlyon & Deeks, 2002; McKay & Carlyon, 1999; van Wieringen et al., 2003).

The discrepancy between single-electrode and multi-electrode pitch sensitivity in CI listeners is thought to be due to interactions between CI electrodes in the cochlea and remains incompletely understood to date. Early stimulation paradigms employed continuous electrical stimuli. With more than one stimulating electrode, these stimuli produced interactions of electric fields due to the broad spread of electric current in the cochlea (Boëx et al., 2003; de Balthasar et al., 2003; Eddington et al., 1978; Favre & Pelizzone, 1993) which severely limited spectral sensitivity. The switch from continuous to pulsatile stimulation, applied sequentially across electrodes, with the introduction of continuous interleaved sampling (CIS, Wilson et al., 1991) avoided electric-field interactions and resulted in greatly improved speech understanding with multi-electrode CIs. However, even with CIS unwanted interactions between electrode signals occur due to across-electrode interactions at the neural level (e.g., de Balthasar et al., 2003), degrading the highly precise neural response to single electric pulses (e.g., Hartmann et al., 1984; Van Den Honert & Stypulkowski, 1987) due to temporal effects such as refractoriness or facilitation (for a review, see Boulet et al., 2016). Such interactions are commonly referred to as channel interactions.

The size and shape of channel interactions crucially depends on both the per-electrode pulse rate(s) and the distribution of pulses (or pulse rates) across electrodes. In particular, the amount of channel interactions should increase with increasing pulse rate and be high for groups of adjacent electrodes with high rates. Yet even most of today's clinically used CI stimulation paradigms employ rather high pulse rates to sufficiently sample the temporal envelope of speech (e.g., Arora et al., 2009; Friesen et al., 2005; Loizou et al., 2000) which has been shown to convey most of the relevant speech information (cf. Smith et al., 2002; but see Zeng et al., 2004). Furthermore, the per-electrode pulse rate is often constant and, hence, does not convey information of the acoustic stimulus. While temporal-envelope cues

predominantly provided by CIS-like stimulation paradigms are sufficient to provide robust speech cues—at least in quiet—and convey gender cues via the fundamental frequency (F0) of voiced speech (roughly in the range of 100 to 250 Hz; e.g., Peterson & Barney, 1952), such cues are insufficient to provide salient temporal cues for pitch (e.g., Gaudrain & Başkent, 2018; Vandali et al., 2005) as well as interaural time differences (ITDs; e.g., Grantham et al., 2008; Laback et al., 2004) outside of laboratory settings.

A major hurdle towards providing more salient temporal cues with multi-electrode stimulation while at the same time maintaining high speech understanding are diverging pulse-rate requirements. While speech is best conveyed at "high" rates beyond 500 pps (e.g., Arora et al., 2009; Friesen et al., 2005), sensitivity to both temporal pitch and ITD worsens when pulse rates increase beyond 200 to 300 pulses per second (pps; e.g., Carlyon et al., 2008; Ihlefeld et al., 2015; Majdak et al., 2006). This rate limitation imposes a low-rate constraint opposite to the high-rate constraint for speech.

Several stimulation paradigms that aim to saliently encode low-rate temporal cues have been proposed. Examples using low rates on all electrodes are fundamental asynchronous stimulus timing (FAST, Smith, 2010) and peak picking (PP, Williges et al., 2018). An example using (rather) high rates on all electrodes is peak-derived timing (PDT, van Hoesel et al., 2008). Furthermore, a family of stimulation paradigms that encode temporal cues via pulse timing on one to four apical electrodes while maintaining CIS stimulation on the other electrodes was introduced (FSP, FS4, and FS4-p; Hochmair et al., 2006; Riss et al., 2014). However, to our knowledge, no published study has yet demonstrated improved temporal coding while fully preserving speech information.

Recently, we developed an approach to improve temporal-cue sensitivity at high pulse rates. It is based on selectively inserting extra pulses shortly after regular pulses, that is, with short inter-pulse intervals (i.e., SIPI pulses). Such stimuli will hereafter be referred to as SIPI

stimuli (cf. Figure 1C). The approach was derived from studies investigating ITD sensitivity with binaurally coherent jitter applied at rates high enough to sufficiently sample the speech envelope (Hancock et al., 2012; Laback & Majdak, 2008). For ITD perception, discrimination performance with unmodulated 1000-pps pulse trains and SIPI pulses inserted at a (SIPI) rate of 100 pps (i.e., after every tenth pulse) was similar to that with unmodulated 100-pps pulse trains (Srinivasan et al., 2018). The latter stimuli will hereafter be referred to as LR ("low rate") stimuli (cf. Figure 1A). SIPI pulses were shown to enhance ITD sensitivity even when imposed on a high-rate pulse trains with low-rate amplitude modulation (AM) (Srinivasan et al., 2020) when the AM depth was less than 100% and the SIPI pulses were inserted at AM peaks (see also Hu et al., 2017). High-rate pulse trains with a superimposed AM will hereafter be referred to as HR stimuli (cf. Figure 1B). Furthermore, Buechel et al. (2018) observed similar effects in a neurophysiological study with unanesthetized rabbits and linked the enhancement with neural facilitation. With respect to temporal pitch, Lindenbeck et al. (2020) showed SIPI-based sensitivity to be at par with LR sensitivity when AM and SIPI rate were consistent and at 125 and 250 Hz irrespective of AM depth. Furthermore, for scenarios in which SIPI and AM rate were inconsistent (i.e., the SIPI rate was an integer sub-multiple of the AM rate as in Srinivasan et al., 2020), Lindenbeck et al. (2023) showed that the SIPIs provide the dominant pitch cue as compared to the AM. Taken together, the single-electrode results gathered so far suggest that the SIPI approach may be promising to enhance temporal-cue sensitivity with high-rate stimuli. However, it has so far only been tested for single-electrode stimulation.

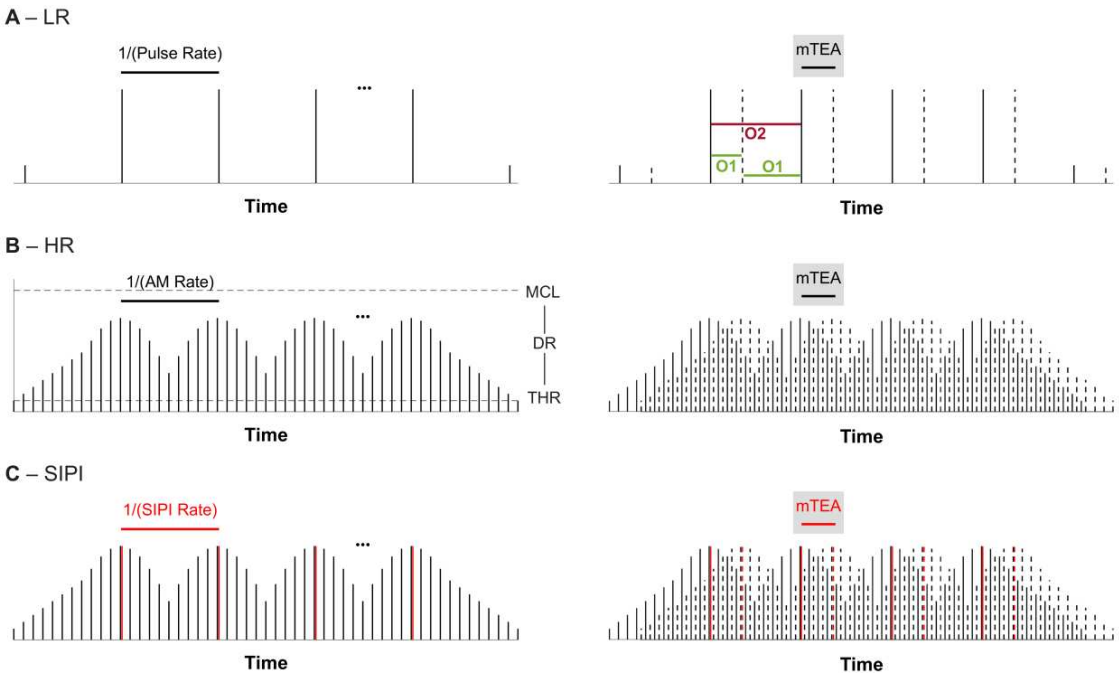


Figure 1

Stimuli used in the experiment. ***Left column****: Single-electrode stimuli (cf. Lindenbeck et al., 2020).* ***Right column****: Dual-electrode stimuli with the monaural temporal electrode asynchrony (mTEA) between the basal (solid line) and the apical (dashed line) electrode.* ***A*** *– LR: Unmodulated pulse train with a rate equal to the F0.* ***B*** *– HR: 1000-pps pulse train with an F0 amplitude modulation (AM) at a depth of 0.3 (Srinivasan et al., 2017). Threshold (THR) and maximum comfortable level (MCL) define the electric dynamic range (DR).* ***C*** *– SIPI: As "B – HR" but with additional pulses inserted with short inter-pulse intervals at every AM peak (thus, the SIPI rate equals the F0). For readability purposes, negative pulse phases are omitted and only a few periods are shown.*

In the current study, we aimed at quantifying the perceptual effects of interactions between two stimulating electrodes on monaural temporal-pitch discrimination for the three stimulus types introduced above. LR stimuli were included to represent the best case for temporal-cue encoding. HR stimuli were included to represent clinical scenarios with CIS-like stimulation. SIPI stimuli were included to assess to what extent the single-electrode SIPI

results (Lindenbeck et al., 2020, 2023; Srinivasan et al., 2018, 2020) generalize to dual-electrode stimulation. The general assumptions were that the (monaural) temporal electrode asynchrony (mTEA; cf. Lindenbeck et al., 2024) between the electrodes determines the degree of channel interactions and, thus, that any effect of mTEA is mediated by the tonotopic separation of the two stimulating electrodes. To alter the potential for channel interactions (i.e., the IPIs; Boulet et al., 2016), we systematically varied the mTEA (see the right column of Figure 1). The mTEA was varied as a proportion of the pulse or AM period, respectively, thus, assuming a periodicity of the mTEA effect. To manipulate the amount of channel interactions (i.e., the summation of the two electrodes within a given auditory-nerve neuron), we systematically varied the tonotopic separation between the component electrodes of the dual-electrode combinations, hereafter referred to as C1 and C2, respectively. C1 was fixed in the center of the electrode array while C2 was varied across the array (cf. Figure 2A) defining various C1-C2 combinations (CCCs). For the large tonotopic separations, the selection was based on a pretest, that is, forward-masked spatial tuning curves determined at C1 (Nelson et al., 2008, 2011). Furthermore, we assumed that the stimulus type *per se* both determines overall sensitivity as well as moderates the mTEA effect. To link our results with the studies using single-electrode stimulation, we furthermore tested single-electrode discrimination sensitivity for all three stimulus types presented at all C1 and C2 electrodes.

In summary, we tested the following hypotheses:

1. For small tonotopic separation of stimulation electrodes (exciting overlapping auditory-nerve fiber populations), sensitivity worsens with increasing mTEA towards 50% of the pulse/AM period and then improves again towards the end of the period.
2. For large tonotopic separations (exciting distinct fiber populations), sensitivity is unaffected by the mTEA.

3. Sensitivity is better with LR and SIPI stimuli as compared to HR stimuli.
4. The mTEA effect is larger for SIPI and HR stimuli (due to their higher overall rates and, hence, higher susceptibility to temporal interactions at the auditory nerve) as compared to LR stimuli.

Methods

Listeners

Five CI listeners (four females) participated in the experiments, see Table 1 for more details. All listeners were post-lingually deafened and used 12-electrode implants manufactured by MED-EL Corp. (Innsbruck, Austria) with electrode spacings of either 2.1 or 2.4 mm, respectively. Bilateral CI listeners were asked to use their preferred ear. We tested only listeners having experience in CI testing in our lab and having at least ten activated electrodes in their clinical fittings. All listeners were volunteers and were paid an hourly wage for their participation.

In the experiments, we followed the European Charter of Fundamental Rights, worked along the guidelines of "Good Scientific Practice", and fulfilled the ethical principles for research involving human subjects (Helsinki declaration). The research protocol was one of the ethics-approved standard protocols for CI research at the Acoustics Research Institute laboratory. All subjects were adults capable of giving informed consent in writing before the experiment. They were adequately informed of the aims, methods, sources of funding, possible conflicts of interest, institutional affiliations of the researchers, the anticipated benefits and potential risks and discomfort it may entail, post-study support, and any other relevant aspects. They were informed of the right to refuse their participation or to withdraw consent to participate at any time without reprisal. When collecting data, we headed for proportionality and avoided collecting more data than necessary. Confidentiality of collected

personal data were maintained by anonymization.

Table 1

Details on the CI listeners. All listeners used 12-electrode MED-EL CIs. If the onset of deafness could not be attributed to a specific date (e.g., in case of progressive hearing loss), "age at onset of deafness" may also refer to the onset of profound hearing loss.

Listener	Sex	Age (yr)	Etiology	Age at onset of deafness (yr)		Ear used for the experiments		
				L	R	Side	CI experience (yr)	Array (implant) type
CI24	female	58	Progressive	40	41	L	12	Standard (C40+)
CI114	female	55	Sudden	40	40	L	15	Standard (Pulsar)
CI116	male	57	Progressive	---	53	R	3	FLEX28 (Synchrony)
CI118	female	39	Sudden	33	---	L	4	FLEX28 (Synchrony)
CI119	female	64	Progressive	61	61	R	2	FLEX28 (Synchrony 2)
Mean ± SD	---	55 ± 9	---	44 ± 12	49 ± 10	---	7 ± 6	---

Stimuli and Apparatus

Figure 1 shows monophasic versions of the three types of biphasic stimuli used in the experiments. We used the parameter F0 to define across stimulus types the rate at which usable temporal-pitch information is conveyed. For LR stimuli (bottom row), the F0 corresponded to the pulse rate. For HR and SIPI stimuli, the F0 corresponded to the AM rate. For SIPI stimuli, the F0 furthermore corresponded to the SIPI rate (i.e., full-rate SIPIs; cf. Lindenbeck et al., 2020, 2023). A SIPI pulse was always inserted 60 µs after and with the same amplitude as the preceding regular pulse. The geometric mean (i.e., nominal) F0 was 100 Hz, however, the actual F0s varied depending on the individual frequency difference

(FD) of the listener. For HR (top row) and SIPI stimuli (middle row), the modulation shape was a full-wave rectified sinusoid and the AM rate *after* rectification corresponded to the F0. The AM was applied proportionally to the dynamic range (DR), that is, the range between absolute threshold (THR) and maximum comfortable level (MCL, cf. top left panel). To ensure quantization of the AM peaks, F0s were integer sub-multiples of the 1000-pps carrier pulse rate used with HR and SIPI stimuli (cf. Lindenbeck et al., 2020; Srinivasan et al., 2020). Note that this discretization also restricted the testable FDs in the F0-discrimination tasks (cf. "Main Experiment: Listener-Specific Frequency Difference" and Lindenbeck et al., 2020). Furthermore, this approach should have prevented beat-frequency distortion (cf. Goldsworthy et al., 2022) and, hence, maximized HR temporal-pitch sensitivity. The modulation depth was 0.3 to simulate the average depth occurring in clinical CI processors (see Srinivasan et al., 2017).

The dual-electrode stimuli (right column of Figure 1) were constructed by presenting the same pulse train at C1 and one of the four C2's. To introduce the mTEA, the more apical stimulus was delayed by up to one AM period (for HR and SIPI stimuli) or one pulse period (for LR stimuli), that is, between 0 and 100%. Six mTEAs were selected: 4, 32, 48, 52, 68, and 96%. These delays were symmetric around 50% and realizable for all stimulus types while avoiding simultaneous stimulation.[4]

Figure 2 shows the four dual-electrode conditions, that is, the four C1-C2 combinations (referred to as CCCs), which are named according to the relative position of C2 (i.e., large-apical [L_A], small-apical [S_A], small-basal [S_B], and large-basal [L_B]). In addition to these four CCCs, a fifth "pseudo-dual" electrode condition (called Z for "zero", in blue) was tested in which both C1 and C2 pulse trains were presented at the C1 electrode, simulating a tonotopic separation of zero (i.e., maximum interaction, cf. McKay & McDermott, 1996). The three CCCs S_A, T, and S_B covered electrode separations from 0 to

2.1/2.4 mm and, hence, the separations of adjacent electrodes for all three major CI manufacturers in terms of market share, that is, Cochlear Inc. (0.75 mm), MED-EL (2.1 or 2.4 mm in our cohort), and Advanced Bionics (1.3 mm; Zeng, 2022). The two largest CCCs L_A and L_B were selected in pretests based on spatial tuning curves (see Section "Forward-Masked Spatial Tuning Curves and C2-Electrode Selection").

All stimuli were nominally 500 ms long, including linear 100-ms onset and offset ramps. The pulse duration (including both phases) was 55 µs, phase duration and inter-phase gap depended on the implant type. The stimuli were generated on a personal computer and were sent directly to the implants via the Research Interface Box II (Institute of Ion Physics and Applied Physics, Leopold-Franzens-University, Innsbruck, Austria).

Loudness Balancing

To avoid loudness confounds on F0 discrimination (e.g., Arnoldner et al., 2008; Vandali et al., 2013), all stimuli used for F0 discrimination were loudness-balanced using the method of adjustment. Provided precautions are taken to avoid starting-level biases (e.g., Van Eeckhoutte et al., 2018) and listeners complying with the task, this method is similarly accurate to adaptive staircase methods (which we have used so far in, e.g., Lindenbeck et al., 2020; Srinivasan et al., 2020) but much faster (about five times in our case, see also Van Eeckhoutte et al., 2018).

Reference and target stimulus were presented using an inter-stimulus break of 400 ms. Per stimulus type, the reference was a single-electrode 100-Hz F0 stimulus presented at the C1 electrode. The listeners were asked to indicate whether both stimuli were equally loud and, if not, which one was louder. They were provided with response buttons changing the amplitude by ± 2, ± 5, or ± 10% of the DR to iteratively adjust the level of the probe stimulus. Per condition, at least two adjustments were performed with the target level starting above the comfortable level for that electrode as determined in the fittings (cf. Section

275 "Electrode Fitting and C1-Electrode Selection") and at least two adjustments with the target
276 level starting below the comfortable level.

277 First, single-electrode stimuli at all C2 electrodes and for all occurring F0s, including
278 those at the C1 electrode, were loudness balanced. Then, using the balanced single-electrode
279 amplitudes, the dual-electrode stimuli with between-electrode delays of 4, 32, and 48% were
280 loudness-balanced by adjusting the amplitudes jointly on both electrodes by the same
281 proportion of the DR (cf. Macherey & Carlyon, 2010). The starting level was roved around
282 ± 20% of the comfortable level in single-electrode conditions and ± 20% of the balanced
283 single-electrode amplitudes in the dual-electrode conditions. Assuming perceptual symmetry
284 along the mTEA range (which was checked) the amplitudes used in the discrimination tests
285 for mTEAs of 52, 68, and 96% were the same as those used for delays of 48, 32, and 4%,
286 respectively.

287 **Electrode and Frequency-Difference Selection Pretests**

288 We conducted a sequence of pretests to find, per CI listener, a suitable set of C1 and
289 C2 electrodes as well as a global FD for all conditions of the main experiment. Listeners were
290 allowed to take breaks at any time.

291 ***Electrode Fitting and C1-Electrode Selection***

292 First, we fitted all electrodes for LR pulse trains by determining THR, MCL (cf.
293 Figure 1) and comfortable level with unmodulated 500-ms 100-pps pulse trains without
294 ramps using an informal adjustment procedure. Based on the fitting, we excluded electrodes
295 inducing uncomfortable sensations and/or having an unexpectedly small DR as compared to
296 the surrounding electrodes.

297 To obtain as many C2 candidate electrodes as possible, we aimed for a C1 electrode in
298 the center of the 12-electrode array (cf. Figure 2A), that is, electrode six was selected for all
299 listeners.

Forward-Masked Spatial Tuning Curves and C2-Electrode Selection

Figure 2A shows the C2 electrode configurations. We aimed at selecting two types of C2 electrodes: C2's with a small tonotopic separation from C1 to induce maximum channel interactions and C2's with a large tonotopic separation from C1 to induce as little channel interactions as possible. As a marker of channel-interaction potential (i.e., across-electrode masking), we measured listener-specific forward-masked spatial tuning curves using the paradigm described in Nelson et al. (2008, 2011).[5]

Before measuring the tuning curves, we fitted all ears for the tuning-curve stimuli using the same procedure as in Section "Electrode Fitting and C1-Electrode Selection". For the masker stimuli, we used unmodulated 160-ms 1000-pps pulse trains without ramps on all available electrodes. For the target stimulus, we used unmodulated 10-ms 1000-pps pulse trains without ramps on the target electrode only.

The tuning curves were measured in a three-interval, three-alternative, forced-choice (3I-3AFC) task. The listeners had to hear out the target that was played randomly in one of the three intervals 10 ms (CI24: 15 ms) after the offset of the masker that was played in all three intervals. Similar to Nelson et al. (2008), the target was presented at the target electrode with a level of 17.0 ± 5.7% (range 10 – 25%) of the target DR. The target level was individually adjusted to avoid floor and ceiling effects due to task difficulty. The probe levels should not have had an effect on the tuning-curve shapes (Nelson et al., 2008). Per masker electrode, the masker level was varied and masking thresholds were determined in an adaptive 2-up 1-down staircase procedure (Levitt, 1971; Leek, 2001) converging at 71% correct probe identification. The initial masker level was 10% of the masker DR. The initial step size was 5% of the masker DR. It was reduced by a factor of 0.7 after each reversal until it reached the minimum step size of 1% of the masker DR. A converging staircase was terminated after 12 reversals. A diverging staircase either exceeded the maximum

comfortable level (i.e., it was aborted before stimulating too loudly) or dropped below the threshold of the masker electrode. The last eight reversals of a converged staircase were averaged to determine the staircase threshold. A masking threshold was calculated as the average of three staircase thresholds.

After having measured the tuning curves, the small-separation C2 electrode pair candidates (one or two neighboring electrodes of C1 on either side) were evaluated for sufficient masking on both ears. For all listeners, the direct neighbors of C1 were selected (cf. Table 2).

Table 2

Setup of the main experiment as determined in the pretests. Per CI listener: The C1 and four C2 electrodes defining the CCCs: large-apical (L_A), small-apical (S_A), zero (Z), small-basal (S_B), and large-basal (L_B), and the corresponding FDs. Statistics for electrode number (El. #) and distance (in mm) from the C1 electrode (defining the Z CCC). Electrodes are numbered from apex to base.

Listener	L_A	S_A	C1/Z	S_B	L_B	FD (%)
CI24	1	5	6	7	9	22
CI114	2	5	6	7	9	22
CI116	2	5	6	7	9	10
CI118	2	5	6	7	9	10
CI119	2	5	6	7	10	10
El. #: Mean ± SD	1.8 ± 0.4	5.0 ± 0.0	6.0 ± 0.0	7.0 ± 0.0	9.2 ± 0.4	---
Distance *re* Z (mm): Mean ± SD	-9.4 ± 1.6	-2.2 ± 0.2	---	2.2 ± 0.2	7.1 ± 0.9	---

The large-separation C2 electrode pairs were selected based on two criteria. The first criterion was the presence of nearly no masking of C1, as indicated by diverging tuning-curve

341 staircases or, at least, largest masked thresholds on the respective side of C1. The second
342 criterion was the tonotopic separation from C1: We estimated the extent of channel
343 interactions by selecting the closest possible non-masking C2. Table 2 summarizes the finally
344 selected large-separation C2 electrodes. On average, large-separation C2's were separated
345 from C1 by about 8 mm on the basal side and about 10 mm on the apical side.
346 After having selected the C2 electrodes, we fitted the C1 and C2 electrodes for the HR
347 and SIPI stimuli with unmodulated 500-ms 1000-pps pulse trains without ramps using the
348 same procedure as described in Section "Electrode Fitting and C1-Electrode Selection".
349 ***Main Experiment: Listener-Specific Frequency Difference***
350 To maximize the number of mTEA and CCC conditions in the main experiment, and
351 to minimize floor and ceiling effects, the listeners were tested at only one listener-specific FD
352 (e.g., Kong et al., 2009; Ihlefeld et al., 2015). Due to the quantization of the AM peak for HR
353 and SIPI stimuli (cf. Section "Stimuli and Apparatus"), only a limited set of F0s was
354 available and, hence, only a limited set of F0 pairs constituting a set of FDs. Consequently,
355 we pre-defined a set of FDs between 10 and 100% whose F0s were geometrically arranged
356 around the nominal F0 of 100 Hz (cf. Kreft et al., 2010; Lindenbeck et al., 2020).
357 To find the listener-specific FDs for the main experiment, an F0-discrimination
358 sensitivity pretest was conducted using LR stimuli (i.e., a rate-discrimination pretest was
359 performed). We used a two-interval, two-alternative, forced-choice (2I-2AFC) task and the
360 method of constant stimuli with several pre-selected FDs depending on prior knowledge on
361 rate-discrimination sensitivity from previous studies in our lab. The stimuli were presented
362 with an inter-stimulus break of 400 ms. Per FD, both possible orders of F0s (low-high and
363 high-low, respectively) were each tested 25 times, but in randomized order. The loudness-
364 balanced amplitudes were roved by 3% of the DR independently per interval. The listeners
365 indicated whether the stimulus in the second interval had a higher or lower pitch than the

stimulus in the first interval. Response feedback was provided after each trial. The collected percent-correct scores were converted to d' scores, accounting for bias towards one order of F0s (Klein, 2001). Per listener, the FD yielding most homogeneous sensitivity across target and flanker electrodes ($d' \geq 1$) while minimizing ceiling effects was selected. Single-electrode LR pitch sensitivity in the pretests was quite homogeneous across listeners. Hence, only two FDs were selected (see Table 2): 10% (91 Hz vs. 100 Hz) and 22% (91 Hz vs. 111 Hz).

In the main experiment, the F0-discrimination task was identical to that used in the pretest. In dual-electrode conditions, level roving was applied similarly on both electrodes. The stimulus types were tested in blocks with the order balanced across listeners. Per stimulus type, to keep place cues constant within a block, the CCCs were tested block-wise. The CCC block order was balanced across listeners but different for each individual listener to (pseudo-)randomize block-order effects. Within each CCC block, the order of mTEAs was randomized. To disentangle interaction effects arising from dual-electrode stimulation from electrode-specific baseline sensitivity, also single-electrode conditions, corresponding to C1 and C2 of all dual-electrode conditions, were measured before dual-electrode conditions within each CCC block. Finally, the entire stimulus type setup described above (i.e., both single- and dual-electrode conditions) was tested twice, each time with 50 repetitions per condition, resulting in 10500 trials in total per listener.

Statistical Analysis

The aim of our statistical analysis was to dissociate both presence and absence of an effect from measurement noise. That is, we aimed to collect evidence for both the null hypothesis (H0, effect absent) and the alternative hypothesis (H1, effect present). However, commonly used null-hypothesis significance testing (NHST) based on p values only insufficiently tests the H0 and has been extensively criticized (Amrhein et al., 2019; Cohen, 1994; Rouder et al., 2016). Thus, to collect H0 evidence but still retain a fixed Type I error

rate and provide estimates of statistical power, we report our statistical results as Bayes-NHST hybrids (cf. Dienes & Mclatchie, 2018; Keysers et al., 2020). Bayesian and NHST results are reported together. For more details regarding this approach, see Lindenbeck et al. (2024).

Bayesian Hypothesis Testing

We used repeated-measures analyses of variance (rmANOVAs) for statistical analysis. In the Bayesian implementation (Keysers et al., 2020; cf. van den Bergh et al., 2020) the Bayes Factor (BF; e.g., Kass & Raftery, 1995) is used to distinguish between *relative* "evidence of presence" (i.e., diagnostic data supporting H1 over H0), "evidence of absence" (i.e., diagnostic data supporting H0 over H1), and "absence of evidence" (i.e., non-diagnostic data that support neither hypothesis). We considered BFs > 3 as evidence for H1 over H0 ("evidence of presence") and used the inverse threshold (i.e., BF < 0.33) to establish evidence of H0 over H1 ("evidence of absence"). BFs between 0.33 and 3 were considered inconclusive ("absence of evidence").

Bayesian statistical analyses were conducted with JASP v0.17.3 (JASP Team, 2023). In particular, we used Bayesian rmANOVAs (Rouder et al., 2012, 2016, 2017; van den Bergh et al., 2020; Van Den Bergh et al., 2023; Wetzels et al., 2012) in the default JASP prior configuration. The output is the so-called BF_{incl} that quantifies evidence for including an effect in the rmANOVA model (H1) over excluding it (H0).

Null-Hypothesis Significance Testing and Confidence Intervals

To provide effect size estimates and well-known Type I error control, we conducted NHST rmANOVAs using MATLAB® R2022b (Mathworks, Inc). The significance level was always 5%. To prevent inflation of Type I errors, we used the Huynh-Feldt correction (Huynh & Feldt, 1970) of *p* values and *F*-test degrees of freedom (cf. Oberfeld & Franke, 2013).

As effect-size metric, we used the generalized η^2 (η_G^2) (Bakeman, 2005; Lakens,

2013; Olejnik & Algina, 2003). It is comparable across within and between-subjects designs and classifies effect sizes as either small ($\eta_G^2 \geq .02$), medium $\eta_G^2 \geq .13$), or large ($\eta_G^2 \geq .26$) (Bakeman, 2005). The so-called exact analytical 90% "within-subjects" confidence intervals (i.e., the 5-to-95% range of η_G^2, the *F*-test is one-sided) are reported in brackets. The confidence interval of a statistically significant effect excludes zero.

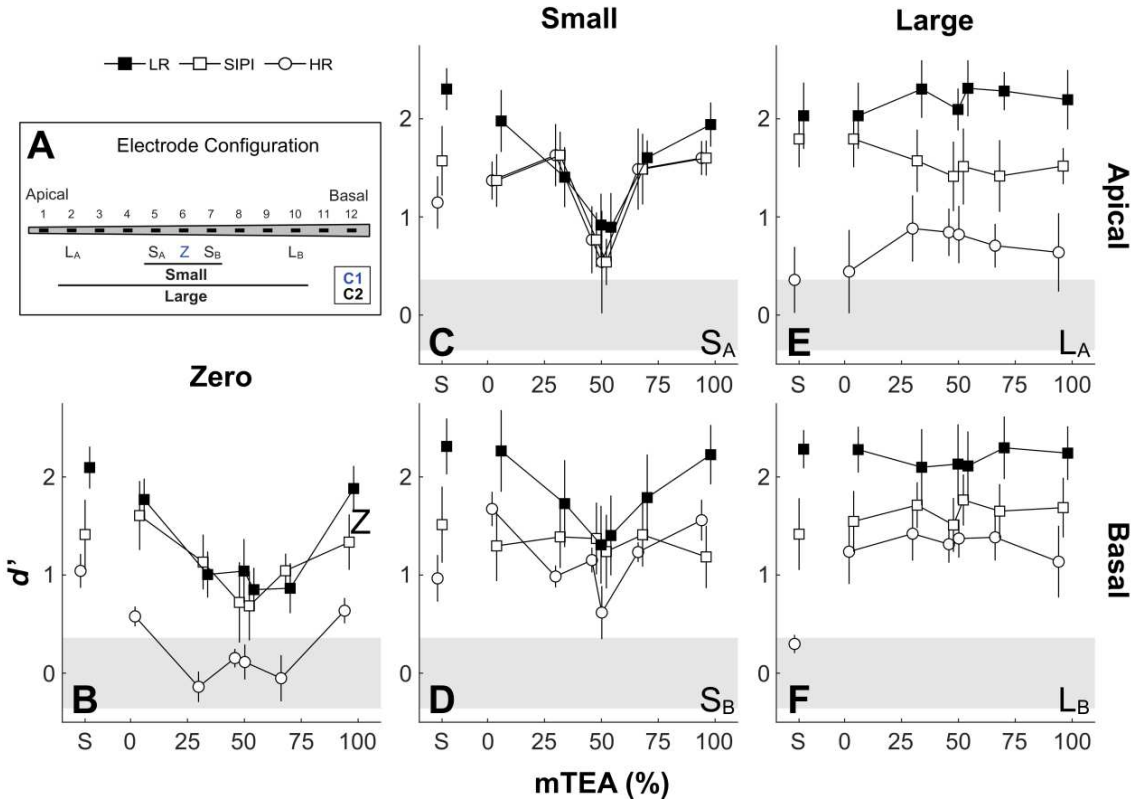


Figure 2

Panel A*: Electrode locations and selection. The C1 electrode was located in the center of the arrangement. There were five options for the C2 electrode, forming five C1-C2 combinations (CCCs). Zero (Z, in blue) used C1 as C2. Two C2 electrodes located on the apical side defined the large-apical (L_A) and small-apical (S_A) CCCs. Two C2 electrodes located on the basal side defined the small-basal (S_B) and large-basal (L_B) CCCs. For S_A and S_B, C2 electrodes were those adjacent to C1. For L_A and L_B, C2 electrodes were selected based on forward-*

masked spatial tuning curves. ***Panels B-F****: F0-based pitch-discrimination sensitivity d' as a function of mTEA (expressed as a proportion of the fundamental period T0 = F0^{-1}) averaged across CI listeners. The left, center, and right column show zero, small, and large tonotopic separation, respectively. Each panel shows a different CCC, cf. panel A). Symbols show the stimulus types (cf. Figure 1). Single-electrode scores for the C2 electrode are shown at "S". The grey area indicates chance performance (|d'| ≤ .36). Error bars show normalized standard errors (Cousineau, 2005; Morey, 2008).*

Results

Single-Electrode Stimulation

In Figure 2B-F, single-electrode listener-averaged d' scores are denoted with an "S" on the abscissa of each panel. Markers distinguish stimulus types. The apex-to-base order of electrodes is L_A-S_A-Z-S_B-L_B (cf. Figure 2A). Figures S1 to S3 provide the individual scores. The LR data are the same as reported in Lindenbeck et al. (2024). The stimulus type seemed to have had an overall effect on sensitivity. HR stimuli yielded the lowest sensitivity, LR stimuli yielded the highest sensitivity, and SIPI-based sensitivity was in between. However, the effect of the stimulus type seemed to have depended somewhat on the electrode.

A two-way rmANOVA with factors Type (LR, SIPI, and HR) and Electrode (L_A, S_A, Z, S_B, and L_B) revealed evidence for a large effect of Type (BF_{incl} = 3.22, F(2.0, 8.0) = 6.04[6], p = .025, η_G^2 = .35 [.03, .55]). The effect of Electrode (BF_{incl} = 0.69, F(2.8, 11.0) = 1.51, p = .266, η_G^2 = .09 [.00, .20]) as well as the interaction between Type and Electrode were inconclusive (BF_{incl} = 2.10, F(8.0, 32.0) = 2.52, p = .030, η_G^2 = .06 [.01, .09]). Figure 3 summarizes the effect of stimulus type across electrodes. Post-hoc pairwise comparisons[7] revealed significantly better LR compared to HR sensitivity (p = .026; other two comparisons p > .159).

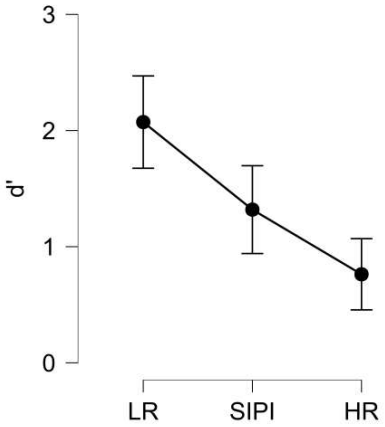


Figure 3

Single-electrode F0-based pitch-discrimination sensitivity d' for different stimulus types averaged across electrodes. Error bars show Bayesian 95% credible intervals (JASP "descriptives plots" output; JASP Team, 2023).

Dual-Electrode Stimulation

Overview

Figure 2B-F shows the dual-electrode *d'* scores as a function of mTEA, with separate panels for each CCC. For each CCC, markers distinguish stimulus types. Figures S1 to S3 provide the individual scores.

First, a clear effect of the mTEA can be observed that is symmetric around 50%. Broadly speaking, low and high mTEAs (short between-electrode delays) show better temporal-pitch sensitivity than intermediate mTEAs (long between-electrode delays). The mTEA effect further seemed to be moderated by the C2 electrode involved (i.e., the CCC). It reduced sensitivity mostly in conditions with no tonotopic separation (i.e., condition Z), a little less in conditions with a small tonotopic separation (i.e., conditions S_A and S_B), and hardly (or even showing slight increases) in conditions with a large tonotopic separation (i.e., conditions L_A and L_B). Second, the stimulus type seemed to have an overall effect like that found for single-electrode stimulation, possibly differing somewhat between CCCs.

To jointly assess the effects of all three experimental parameters, we conducted a three-way rmANOVA with the factors mTEA (4%, 32%, 48%, 52%, 68%, and 96%), CCC (LA, SA, Z, SB, and LB), and Type (LR, HR, and SIPI). We found evidence for a small interaction between mTEA and CCC (BF_{incl} = 126346.17, $F(4.9, 19.7) = 5.30$, $p = .003$, $\eta_G^2 = .07$ [.02, .10]) as well as evidence for the absence of an interaction between mTEA and Type (BF_{incl} = 0.09, $F(5.1, 20.4) = 1.59$, $p = .207$, $\eta_G^2 = .01$ [.00, .02]). The other interactions were inconclusive (see Table S1). Furthermore, we found evidence for a small effect of mTEA (BF_{incl} = 13.12, $F(2.5, 10.1) = 5.80$, $p = .017$, $\eta_G^2 = .06$ [.01, .10]) and a medium effect of CCC (BF_{incl} = 42.16, $F(3.4, 13.4) = 7.96$, $p = .002$, $\eta_G^2 = .16$ [.06, .24]). The evidence for a Type effect was inconclusive (see Table S1).

To further assess the interaction between mTEA and CCC, we investigated the mTEA effect separately for different tonotopic separations.

Large Tonotopic Separation

For the CCC's L_A and L_B, we hypothesized the *absence* of an mTEA effect. To this end, we conducted an rmANOVA with the effects mTEA and Type as specified above as well as the effect of CCC (L_A and L_B only). The mTEA effect depended neither on CCC (absence of evidence for the mTEA × CCC interaction; BF_{incl} = 0.08, $F(2.7, 10.8) = 0.28$, $p = .819$, $\eta_G^2 = .00$ [.00, .00]) nor Type (absence of evidence for the mTEA × Type interaction; BF_{incl} = 0.16, $F(4.7, 18.9) = 1.49$, $p = .243$, $\eta_G^2 = .01$ [.00, .02]). In fact, the mTEA effect was absent altogether (BF_{incl} = 0.07, $F(5.0, 20.0) = 0.75$, $p = .596$, $\eta_G^2 = .00$ [.00, .01]). All other main effects and interactions were inconclusive (see Table S2).

In the right panel of Figure 4, the (absent) mTEA effects are visualized separately per CCC and based on simple-effects rmANOVAs with the factors mTEA and Type (because the main effect of Type and many interactions including it were inconclusive, see Table S2). Table S3 summarizes their results.

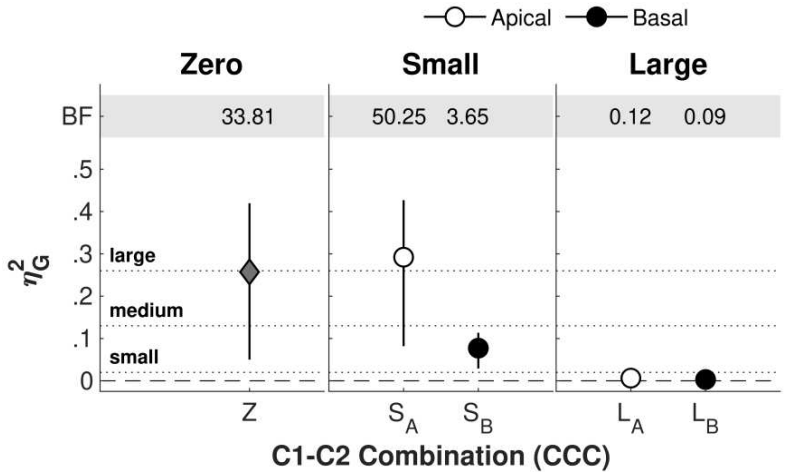


Figure 4

Effect-size estimates η_G^2 with 90% confidence intervals for the three tonotopic separations (panels) and corresponding CCCs (abscissa, cf. Figure 2A). The dashed line indicates an η_G^2 of zero; effects whose confidence intervals do not touch this line are significant with $p < .05$. Dotted lines indicate threshold η_G^2 for small, medium, and large effects, respectively. Bayes factors (BFs) in the grey ribbon indicate evidence for the presence of an mTEA effect (BF > 3) re evidence of its absence (BF < 0.33), cf. Section "Statistical Analysis".

Zero and Small Tonotopic Separations

For the CCC's Z, S_A, and S_B, we hypothesized the *presence* of an mTEA effect. To this end, we conducted an rmANOVA with the effects mTEA and Type as specified above as well as the effect of CCC (Z, S_A, and S_B only). The mTEA effect still depended on the CCC (BF_{incl} = 7.84, $F(10.0, 40.0) = 4.26$, $p < .001$, $\eta_G^2 = .04$ [.02, .05]) and there was evidence for medium effects of mTEA (BF_{incl} = 69.15, $F(1.8, 7.3) = 7.31$, $p = .019$, $\eta_G^2 = .17$ [.04, .29]) and CCC (BF_{incl} = 4.87, $F(1.3, 5.3) = 7.55$, $p = .034$, $\eta_G^2 = .12$ [.02, .23]). All other effects were inconclusive (see Table S4). Importantly, this means that for CCCs showing an mTEA effect, the effect of the stimulus type is somewhat spurious and not *per se* absent as for the CCCs not showing an mTEA effect.

To further assess this, we conducted simple-effects rmANOVAs with the factors mTEA and Type separately per CCC. The mTEA effects across Types (i.e., main effects of mTEA) are visualized in the left and middle panel of Figure 4 and summarized in Table S3. An mTEA effect is present for all CCCs, however, larger for the Z and S_A CCC's than for the S_B CCC. Furthermore, for the S_B CCC, the mTEA effect differed between Types (BF_{incl} = 5.08, $F(10.0, 40.0) = 2.18$, $p = .040$[8], $\eta_G^2 = .06$ [.00, .09]). For the other two CCCs, the effect of Type on the mTEA effect was inconclusive, yet, close to being absent[9] (Z: BF_{incl} = 0.35, $F(9.9, 39.7) = 1.11$, $p = .379$, $\eta_G^2 = .06$ [.00, .07]; S_A: BF_{incl} = 0.38, $F(10.0, 40.0) = 1.38$, $p = .226$, $\eta_G^2 = .03$ [.00, .04]). For the S_B CCC, only LR (BF_{incl} = 3.66, $F(5.0, 20.0) = 4.12$[10], $p = .010$, $\eta_G^2 = .21$ [.01, .35]) and HR (BF_{incl} = 7.40, $F(2.8, 11.3) = 4.30$, $p = .032$, $\eta_G^2 = .30$ [.04, .44]) stimuli showed an mTEA effect. For SIPI stimuli, it was absent (BF_{incl} = 0.27, $F(3.2, 13.0) = 0.74$, $p = .555$, $\eta_G^2 = .01$ [.00, .02]).

Interim Summary

So far, we have shown that for single-electrode stimulation the stimulus type has a substantial effect on pitch-discrimination sensitivity, with LR stimuli yielding highest and HR stimuli yielding lowest sensitivity, and that the electrode does not have a conclusive effect. Furthermore, for dual-electrode stimulation, we have shown that the effect of the mTEA (quantified as the effect size across all tested mTEAs) is substantial and present for zero- and small-separation CCCs and absent for large-separation CCCs. Both the presence and the absence of the mTEA effect are largely similar across stimulus types with the only exception being the S_B CCC for which LR and HR stimuli showed an mTEA effect while SIPI stimuli did not.

To better understand the origin of the mTEA effect for small (including zero) tonotopic separations, we next assess pitch-discrimination sensitivity separately for certain mTEAs by comparison with single-electrode stimulation.

Table 3

Post-hoc t-tests conducted separately for LR and HR stimuli to assess differences between single-/dual-electrode stimulation (factor Condition), separately for zero and small tonotopic separations, as well as differences between separations for a certain stimulation condition. |t|…test statistic (absolute value); df…error degrees of freedom from the rmANOVA interaction; p_{BH}…two-sided Benjamini-Hochberg-corrected[11] significance level (Benjamini & Hochberg, 1995); d…effect-size estimate [95% confidence interval]. Significant comparisons are marked with an asterisk (*).

Reference	Comparison	LR				HR			
		|t|	*df*	p_{BH}	*d [2.5%, 97.5%]*	*|t|*	*df*	p_{BH}	*d [2.5%, 97.5%]*
Zero, S	Zero, 4/96%	2.17	4	.144	1.1 [0.3, 1.8]	1.63	4	.177	0.7 [0.0, 1.5]
	Zero, 32/68%	5.45	4	.013*	2.7 [2.0, 3.5]	4.28	4	.029*	1.9 [1.2, 2.6]
	Zero, 48/52%	5.41	4	.013*	2.7 [1.9, 3.4]	3.42	4	.049*	1.5 [0.8, 2.3]
Small, S	Small, 4/96%	0.05	4	.241	0.0 [-0.7, 0.8]	1.00	4	.237	0.4 [-0.3, 1.2]
	Small, 32/68%	1.78	4	.187	0.9 [0.1, 1.6]	0.32	4	.237	0.1 [-0.6, 0.9]
	Small, 48/52%	3.62	4	.045*	1.8 [1.0, 2.6]	1.97	4	.165	0.9 [0.1, 1.6]
Zero, S	Small, S	1.33	4	.215	0.5 [-0.1, 1.2]	0.07	4	.237	0.0 [-0.6, 0.6]
4/96%, Zero	4/96%, Small	1.23	4	.215	0.5 [-0.1, 1.1]	3.37	4	.049*	1.2 [0.6, 1.8]
32/68%, Zero	32/68%, Small	3.08	4	.064	1.3 [0.6, 1.9]	5.03	4	.018*	1.8 [1.2, 2.4]
48/52%; Zero	48/52%, Small	0.82	4	.229	0.3 [-0.3, 1.0]	1.89	4	.165	0.7 [0.1, 1.3]

Single-Electrode vs Dual-Electrode Stimulation for Small Tonotopic Separations

To assess dual-electrode stimulation for certain mTEAs in relation to single-electrode stimulation, we focused on small and zero tonotopic separation (i.e., CCC's Z, S_A, and S_B), and stimulus types which showed an mTEA effect for all three relevant CCC's before (i.e., LR and HR). We conducted a three-way rmANOVA with factors Condition (S [single-

554 electrode], 4/96% [= 4% and 96% averaged due to symmetry], 32/68%, and 48/52%),
555 Separation (zero and small [= S_A and S_B averaged]), and Type (LR and HR).
556 There was evidence for the absence of a three-way interaction (BF_{incl} = 0.25,
557 $F(1.5, 6.2) = 0.10$, $p = .858$, $\eta_G^2 = .00$ [.00, .02]). However, there was evidence for a small
558 interaction between Condition and Separation (BF_{incl} = 72.24, $F(2.6, 10.5) = 36.58$, $p < .001$,
559 $\eta_G^2 = .11$ [.08, .15]). Furthermore, there was evidence for a large effect of Condition
560 (BF_{incl} = 50.11, $F(1.6, 6.3) = 11.08$, $p = .010$, $\eta_G^2 = .38$ [.13, .54]), a small effect of Separation
561 (BF_{incl} = 3.52, $F(1.0, 4.0) = 27.60$, $p = .006$, $\eta_G^2 = .12$ [.06, .19]), and a large effect of Type
562 (BF_{incl} = 8.05, $F(1.0, 4.0) = 20.65$, $p = .010$, $\eta_G^2 = .46$ [.20, .61]). All other interactions were
563 inconclusive (see Table S5).
564 Figure 5 visualizes the interaction between Condition and Separation as well as the
565 main effect of Type (different panels). Tables S6 and S7 summarize the results of the
566 corresponding type-specific rmANOVAs. To assess these interactions further, we conducted
567 post-hoc *t*-tests[12] on the type-specific rmANOVAs, the results of which are summarized in
568 Table 3. For both stimulus types, in the zero-separation condition, dual-electrode stimulation
569 yielded lowered sensitivity for mTEAs of 32/68% and 48/52% but not for 4/96%, as
570 compared to single-electrode stimulation (condition S). For LR stimuli, in the small-
571 separation condition, dual-electrode stimulation yielded lowered sensitivity only for mTEAs
572 of 48/52%. Furthermore, for mTEAs of 32/68%, sensitivity was lower in the zero-separation
573 condition as compared to the small-separation condition – yet, for LR stimuli, only
574 approaching significance.
575 In summary, dual-electrode stimulation never yields significantly higher sensitivity
576 than single-electrode stimulation. Instead, dual-electrode sensitivity worsens with increasing
577 mTEA (symmetrically towards 50%) as compared to single-electrode sensitivity. The
578 tonotopic separation has an effect for intermediate mTEAs (i.e., 32/68%).

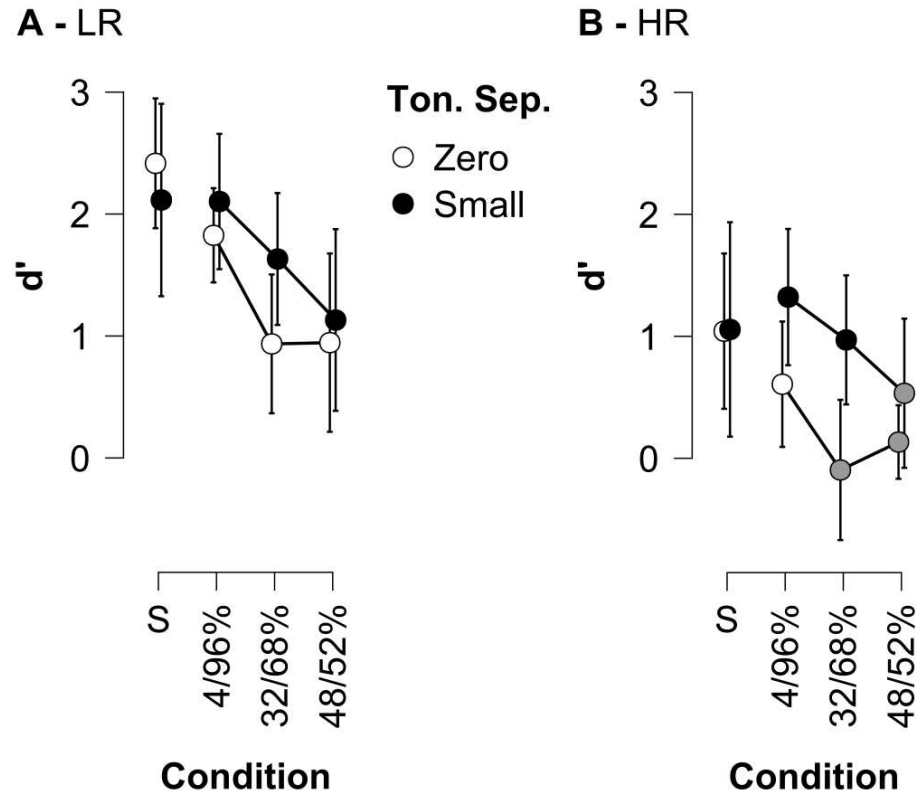


Figure 5

*Pitch-discrimination sensitivity d' for single- ("S") and dual-electrode stimulation for zero and small tonotopic separations, separately for LR (panel **A**) and HR (panel **B**) stimuli. Error bars show Bayesian 95% credible intervals (JASP "descriptives plots" output; JASP Team, 2023). Grey filling of a point indicates that, based on bootstrapped marginal means (1000 repetitions), the average d' is not significantly different from zero (i.e., $p > .05$) and, hence, absent sensitivity at the group level.*

Discussion

Effect of the monaural temporal electrode asynchrony

We hypothesized that the mTEA influences temporal-pitch sensitivity for small tonotopic electrode separations with at least partly overlapping neural populations. To find such separations, we measured forward-masked spatial tuning curves for the C1 electrode and

searched for the nearest neighboring electrodes (C2) that substantially masked the C1 electrode. The resulting small-separation conditions S_A and S_B were separated by on average 2.2 mm. A pseudo-dual-electrode condition Z with a separation of 0 mm was furthermore tested to compensate for the rather large electrode separation of MED-EL CIs. Based on the tuning curve measurements, we also identified C2 electrodes that did not (or at least only to a small amount) mask the C1 electrode and selected large-separation conditions L_A (9.4 mm on average) and L_B (7.1 mm on average) to test the hypothesis that any mTEA effect vanishes for sufficiently large separations. Finally, the size of mTEA effects was hypothesized to depend on the stimulus type, that is, the number of pulses and pulse density. Therefore, low-rate (i.e., LR) stimuli were assumed to yield smaller mTEA effects than high-rate (HR and SIPI) stimuli.

Consistent with the spatial-separation hypothesis, we found large mTEA effects for all zero- and small-separation conditions and the absence of mTEA effects for large-separation conditions. However, mTEA effects for the zero- and small-separation conditions were found not only for the HR but also for the LR condition. This indicates that the mTEA effect does not simply depend on the pulse number or density, respectively (see also Section “Low-Rate vs. High-Rate Stimuli”).

Several previous studies have investigated mTEA effects with loudness-balanced stimuli for a subset of our experimental conditions (Egger et al., 2016; Fielden et al., 2014; Francart et al., 2015; Griessner et al., 2021; Lindenbeck et al., 2024; Macherey & Carlyon, 2010; McKay & McDermott, 1996). With dual-electrode HR stimuli and several mTEAs, McKay & McDermott (1996) showed that with tonotopic separations between 0.75 and 3.75 mm CI listeners perceive the aggregate pulse pattern instead of the per-electrode patterns for separations of 3 mm and less. Also for HR stimuli but triple-electrode stimulation with either 0.75 or 3 mm tonotopic separation, Francart et al. (2015) found a detrimental mTEA

effect on ITD sensitivity for broadly similar mTEAs as in the present study and particularly for the 0.75-mm separation. Both HR studies are consistent with the results we found for HR stimuli, showing detrimental mTEA effects for tonotopic separations of about 3 mm or less.

With dual-electrode LR stimuli and using a ranking task, both Macherey & Carlyon (2010) and Griessner et al. (2021) found little to no mTEA effect on the pitch rank for tonotopic separations of 0.75-1.1 mm and 2.1 mm, respectively. Yet, using dual-electrode LR stimuli and a discrimination task, both Egger et al. (2016) and Lindenbeck et al. (2024) found detrimental mTEA effects on ITD sensitivity for tonotopic separations less than 6 and 3 mm, respectively. In particular, Lindenbeck et al. (2024) tested a (largely) different CI listener cohort but with a matched experimental setup to the present study. They directly compared ITD and temporal-pitch sensitivity and showed that, at least for LR stimuli, the mTEA effect did not differ between temporal-pitch and ITD perception. Furthermore, Fielden et al. (2014) tested mTEA discrimination with dual-electrode LR stimuli and found detrimental mTEA effects for tonotopic separations of about 2.1 mm and less.

The discrepancy in the outcomes are most likely to origin from the psychophysical task, that is, the employment of a discrimination task (strong mTEA effects for tonotopic separations between 2 to 6 mm or less; Egger et al., 2016; Fielden et al., 2014; Lindenbeck et al., 2024) instead of a ranking task (little to no effects; Griessner et al., 2021; Macherey & Carlyon, 2010). Neurophysiological data by Middlebrooks (2004), recorded from the auditory cortices of anesthetized guinea pigs, are more in line with the psychophysical LR pitch-ranking data. In that study, for dual-electrode 254-pps LR stimuli an mTEA of ~50% did not yield channel interactions (i.e., a change in neuronal thresholds *re* single-electrode stimulation). Still, for per-electrode pulse rates above 1000 pps, channel interactions started to increase, broadly in line with the present HR results as well as the previous results collected using both ranking (McKay & McDermott, 1996) and discrimination tasks (Francart

et al., 2015). Hence, we speculate that, for low pulse rates, ranking and discrimination tasks may foster different central analysis strategies. For instance, it may be the case that in ranking tasks electrode signals are more likely analyzed separately while in discrimination tasks they are integrated and analyzed jointly. Furthermore, we note that the rate differences to be ranked in Macherey & Carlyon (2010) and Griessner et al. (2021) were generally at least twice as large as in the present study. It might be the case that mTEA effects manifest differently close to threshold as compared to sufficiently above threshold.

Beyond the psychophysical task, other experimental parameters such as the CI manufacturer and associated electrode array design might potentially play a role. However, neither of the factors electrode array (Cochlear: Francart et al., 2015; Macherey & Carlyon, 2010; McKay & McDermott, 1996; MED-EL: Egger et al., 2016; Griessner et al., 2021; Lindenbeck et al., 2024; Advanced Bionics: Fielden et al., 2014), stimulation mode (monopolar: Egger et al., 2016; Fielden et al., 2014; Francart et al., 2015; Griessner et al., 2021; Lindenbeck et al., 2024; bipolar: Macherey & Carlyon, 2010; McKay & McDermott, 1996), or tested cue (temporal pitch: Griessner et al., 2021; Macherey & Carlyon, 2010; McKay & McDermott, 1996; ITD: Egger et al., 2016; Francart et al., 2015; Lindenbeck et al., 2024; any cue [oddball task]: Fielden et al., 2014) varies consistently with the reported mTEA effects. These factors are unlikely to underlie the differences in LR mTEA effects. In any case, it is evident that more research is necessary to better understand the discrepancy between studies.

In addition to LR and HR stimuli, we also tested dual-electrode SIPI stimuli. This has not been done before and, hence, there is no previous research available for comparison. In theory, on the one hand, SIPI stimuli should produce at least as much channel interaction as HR stimuli as they contain even more pulses (cf. Figure 1). On the other hand, temporal pitch coding should be enhanced, as single-electrode SIPI stimulation has previously been shown

to be at par with LR stimulation (Buechel et al., 2018; Lindenbeck et al., 2024; Srinivasan et al., 2018, 2020). Since LR and HR stimulation yielded similar mTEA effects, it comes as no surprise that SIPI stimuli showed largely similar mTEA effects. However, there is one exception, that is, the absent mTEA effect for the S_B condition for which we currently cannot provide an explanation. We stress that the effect was truly absent (in the Bayesian sense) rather than being inconclusive. Hence, this finding was not due to insufficient power. Furthermore, this absence was almost fully consistent among CI listeners (cf. Figure S2). Hence, it may be the case that this outlier in our study is a result of a particular combination of mTEA, tonotopic separation, and stimulus type (hence, IPIs). Future research may investigate if and in which configurations this generally beneficial effect can be generalized to other multi-electrode setups. If such a generalization were possible, the SIPI approach may in fact be a promising candidate to improve temporal sensitivity in multi-electrode high-rate stimulation.

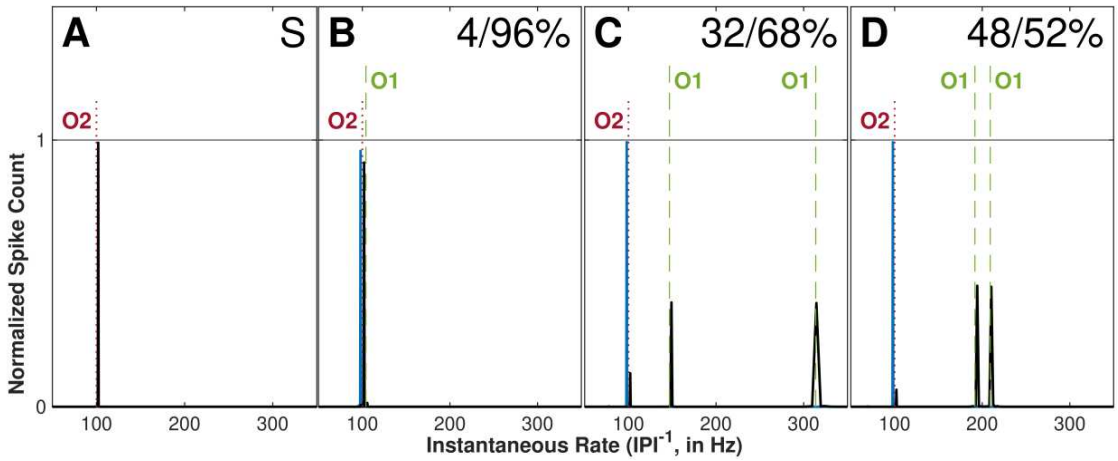
Low-Rate Stimuli: Tonotopic Separation, Effective Rate, and Temporal-Pitch Theories

We hypothesized that the tonotopic separation determines the amount of channel interactions, from the perspective of an auditory-nerve neuron proximal to the C1 electrode, imposing a form of weighting of the C2 pulse train depending on the tonotopic distance and, hence, determining the compound pulse pattern for each individual nerve fiber. Such a weighting should have been 1 for the Z CCC, 0 for the L_A and L_B CCC, and somewhere in between (and probably closer to 1 than to 0) for the S_A and S_B CCCs. In case of HR and SIPI stimulation, the tonotopic separation adds another AM, on top of that already included in the per-electrode stimuli.

Our results for LR stimuli with zero or small separation are particularly well suited to re-evaluate theories of IPI-based temporal pitch (Carlyon et al., 2002; van Wieringen et al., 2003) and discuss the reported LR mTEA effects in the light of the perceptual evaluation of

the neural IPIs. In detail, our zero-separation LR condition closely resembles the unmodulated irregular-rate stimuli evaluated in Carlyon et al. (2002) while our small-separation LR conditions resemble the amplitude-modulated irregular-rate stimuli evaluated in van Wieringen et al. (2003). Irregular-rate stimuli are characterized by two types of IPIs (cf. right column of Figure 1A). First-order IPIs (O1, in green) describe the time interval between successive pulses. O1s alternate between two values: in Carlyon et al. (2002), they mainly alternated between 4 and 6 ms. In our study, the O1s are determined by the mTEA (either mTEA / 100 * T0 or (100 – mTEA) / 100 * T0, $T0 = F0^{-1}$). For example, for the mTEA of 32%, the resulting O1s were 3.2 and 6.8 ms (T0 = 10 ms). Second-order IPIs (O2, in red) describe the time interval between one and the next-but-one pulse. Even with irregular stimuli, O2s are regular and, both in Carlyon et al. (2002) and the present study, corresponded to 10 ms, that is, the T0.

In Carlyon et al. (2002), the perceived pitch of any of these irregular-rate stimuli was never that of the O2 as would be predicted by autocorrelation but a weighted sum of the O1s with much more weight given to the larger O1. Hence, the perceived pitch of a 4-6 irregular-rate pulse train would closely match that of a regular pulse train with 6 ms IPI, that is, 167 pps (in fact, it would match 176 pps). A similar conclusion was reached by McKay & McDermott (1996) for conditions in which their listeners perceived the aggregate dual-electrode pulse pattern. Furthermore, the additionally imposed AM in van Wieringen et al. (2003)—which can be thought of as reducing the amplitudes of the dashed pulses in the right column of Figure 1A and thereby representing the tonotopic separation in our study (i.e., the increasing distance of the C2 electrode to the C1 electrode)—reduced the perceived pitch from the larger O1 towards the O2, that is, towards the T0 and thereby aligning IPI-based and autocorrelation-based temporal-pitch theories.



718

719 **Figure 6**

720 *Spike counts as a function of instantaneous rate (i.e., IPI^{-1}, in Hz) for **LR** stimuli. To remove*
721 *the main effect of stimulus type (cf. Figure 3 and Figure 9), counts are normalized to the total*
722 *number of spikes in a certain condition. Panel **A**: single-electrode stimulation. Panels **B** - **D**:*
723 *dual-electrode stimulation. The mTEAs are denoted in the top right of each panel. Zero-sepa-*
724 *ration conditions in black, small-separation conditions (2.2 mm, cf. S_A and SB in Table 2) in*
725 *blue. Rates corresponding to first-order (O1) and second-order (O2) IPIs (cf. Carlyon et al.,*
726 *2002; van Wieringen et al., 2003) are denoted with vertical dashed green and dotted red*
727 *lines, respectively. Table 4 summarizes the weighted average instantaneous rates.*

728 To *qualitatively* assess the validity of this IPI-based hypothesis on the dominant rate
729 with LR stimuli with varying mTEAs—or varying irregularity—, we simulated AN spiking
730 patterns in response to our LR stimuli with an F0 of 100 Hz using the phenomenological
731 sequential biphasic leaky integrate-and-fire (S-BLIF) model (Horne et al., 2016; Takanen &
732 Seeber, 2022). The model assumes an AN fiber being proximal to a C1 electrode. The
733 parameter set of the original model optimized for animal data (Takanen & Seeber, 2022) was
734 slightly modified so that the model produces thresholds of the electrically-evoked compound
735 action potential (eCAP) in the human range (cf. Takanen et al., 2024). For further details on
736 the modeling rationale as well as model parameterization and validation, see the Appendix.

Figure 6 shows the converted model output, that is, the normalized spike counts as a function of the instantaneous rate (i.e., the IPI^{-1}), for LR stimuli. In the following, the model output will be compared to the psychophysical data from Figure 5A, in particular, the reduction in dual-electrode temporal-pitch sensitivity for mTEAs of 32/68 and 48/52% for the zero-separation condition relative to single-electrode sensitivity as well as the trend towards increased dual-electrode sensitivity for the small-separation condition compared to the zero-separation condition for mTEA of 32/68%.

Figure 6A shows the single-electrode condition. The O2, that is, the T0, is almost perfectly coded. In Figure 6B-C the black lines denote the zero-separation condition. For the 4/96% condition (panel B) again the O2 is almost perfectly coded. This is the case because the 4% mTEA corresponds to an IPI of 400 µs which is well within the range of refractoriness (see, e.g., Boulet et al., 2016). This neural similarity to the single-electrode condition is also consistent with the similar *d'*s for these conditions (cf. Figure 5A).

For the 32/68% condition (panel C), the O1s are coded equally well and the O2 is coded less frequently. The general dominance of the O1s is well in line with the results by Carlyon et al. (2002), however, the equal prominence of the O1s suggests that the dominance of the larger O1 (corresponding to the *lower* instantaneous rate) reported by Carlyon et al. (2002) may be induced by processing stages central to the AN. Implicitly, the equal prominence of the O1 corresponding to an instantaneous rate of 313 pps also suggests that rate-limitation effects reported to lower perceptual sensitivity already for rates below 300 pps (e.g., Carlyon et al., 2008; Ihlefeld et al., 2015) are either not captured by the AN model and/or manifest central to the AN (e.g., Carlyon et al., 2010). The equal prominence, however, would be well in line with studies reporting higher rate limits than 300 pps, particularly for MED-EL CIs (De Groote et al., 2024; Kong et al., 2009).

Table 4

Spike-count-weighted mean (M) instantaneous rate (in Hz; i.e., weighted averaged IPI-1, cf. panels B - D of Figure 6, Figure 7, and Figure S6) alongside spike-count-weighted standard deviation (SD) for dual-electrode simulation conditions: different stimulus types and mTEAs (cf. Figure 1, right column). Frequency changes ΔF as linear differences to reference condition (in Hz). S… single-electrode stimulation (cf. Table 5). For the actual currents used, see Table 6.

Type	mTEA	Zero separation		Small separation							
		Experimental setup		*Experimental setup*		*0.75 mm Separation*		*Amps increased to Z level*		*Small-Zero Amp. Difference doubled*	
		M (SD)	ΔF *re* S	M (SD)	ΔF *re* S	M (SD)	ΔF *re* S	M (SD)	ΔF *re* S	M (SD)	ΔF *re* S
LR	4/96%	100 (1)	0	100 (0)	0	100 (1)	0	100 (0)	0	100 (0)	0
	32/68%	211 (89)	111	100 (7)	0	159 (86)	59	101 (11)	1	100 (0)	0
	48/52%	194 (26)	94	100 (0)	0	151 (51)	51	100 (2)	0	100 (0)	0
HR	4/96%	162 (76)	33	130 (61)	1	127 (59)	-2	154 (75)	25	114 (49)	-15
	32/68%	197 (81)	68	182 (84)	53	168 (79)	39	188 (86)	58	170 (80)	41
	48/52%	195 (82)	66	169 (81)	40	161 (77)	31	181 (85)	52	151 (75)	22
SIPI	4/96%	135 (70)	21	112 (42)	-3	108 (35)	-7	129 (62)	14	102 (17)	-13
	32/68%	209 (74)	94	148 (74)	34	155 (75)	40	173 (84)	59	126 (58)	11
	48/52%	206 (56)	91	136 (67)	22	155 (65)	40	161 (81)	46	120 (53)	5

For the 48/52% condition (panel D), the prominence of O1s and O2s is similar to the 32/68% condition. Considering that for these two conditions the *d'*s dropped by approximately the same amount for the dual- compared to the single-electrode configurations, the interpretation of the model response requires some additional assumptions to be compatible with the psychophysical results pattern. The first scenario would be a perceptual

dominance of the left-most O1 (48/52%: 192 pps; 32/68%: 147 pps) as in Carlyon et al. (2002) in combination with a rate limitation reducing perceptual sensitivity already for dominant rates of about 150 pps. The second scenario would be a perceptual dominance of the shorter (right-most) O1 (48/52%: 209 pps; 32/68%: 314 pps) (cf. the high-voice superiority effect in music; Trainor et al., 2014) in combination with a rate limitation (e.g., De Groote et al., 2024) reducing perceptual sensitivity for dominant rates exceeding about 300 pps. The latter scenario, however, should result in no decline in sensitivity at all. Hence, in light of the model output, the combined results for the 48/52% and 32/68% suggest either a rate limitation reducing sensitivity already at rates of about 150 pps at stages central to the AN or another non-pitch aspect not considered so far. Future research may further clarify the origin of the decline in perceptual sensitivity with increasing rates and heterogeneous IPIs.

In Figure 6B-D, blue lines depict the model output for the small-separation conditions. For all dual-electrode conditions, the model predicts full independence of electrodes, in line with neurophysiological data (e.g., Middlebrooks, 2004), but contrary to the lower *d'*s found for the 32/68% and 48/52% conditions. Simulations with smaller tonotopic separation or small-separation amplitudes, explained in detail in the Appendix, revealed that reducing the separation produced a model output more in line with the psychophysical results. Changes in amplitude had little effect on the model output. In summary, the results suggest that the model underestimates the spread of excitation. Future research and modeling work may clarify the exact effects of changes in tonotopic separation.

Low-Rate vs. High-Rate Stimuli

We hypothesized that the type of stimulus (low- or high-rate) determines overall dual-electrode sensitivity, broadly similar to single-electrode stimulation. Moreover, LR and SIPI stimuli were assumed to yield similarly high sensitivity and HR stimuli (for the selected AM depth of 0.3) were assumed to yield worse sensitivity (cf. Lindenbeck et al., 2020). Indeed,

our single-electrode data showed a large effect of stimulus type and LR sensitivity to be significantly higher than HR sensitivity across all five tested C1 and C2 electrodes. Yet, SIPI sensitivity was neither significantly better than HR sensitivity nor significantly worse than LR sensitivity. Compared to Lindenbeck et al. (2020), in absolute terms LR sensitivity was broadly similar while both HR and SIPI sensitivity was lower in the present study. Furthermore, the difference between HR and SIPI sensitivity was similar in the two studies. Taken together, this suggests a larger variance in the present study (note that both studies had the same sample size) combined with a lower base sensitivity to high-rate stimulation in the CI listener cohort of the present study.

For dual-electrode stimulation, the stimulus type had a large overall effect on sensitivity similar to single-electrode stimulation, however, close to no moderating effect on the mTEA effect (see Section “Effect of the monaural temporal electrode asynchrony” for in-depth discussion). Hence, taken together, the averaged dual- and single-electrode results reflect previous results showing better temporal-pitch sensitivity to pulse rate as compared to AM rate (e.g., Goldsworthy et al., 2021, 2022). Furthermore, dual-electrode sensitivity was never better than single-electrode sensitivity, well in line with previous studies on pitch and ITD perception with loudness-balanced stimuli (e.g., Egger et al., 2016; Galvin et al., 2015). Still, other statistical results for dual-electrode stimulation and involving the stimulus type were largely inconclusive which, in combination with the large effect sizes, suggests that the stimulus type effect was highly variable across listeners. This appears to highlight the need for individualization in the stimulation approach with CIs.

Above, we have already qualitatively discussed the potential neural basis for dual-electrode low-rate temporal-pitch sensitivity. Towards better understanding the neural bases also with physically more complex (but clinically relevant) high-rate AM stimuli, we used the same phenomenological AN model with our HR stimuli as input. Figure 7 shows the model in

the same format as in Figure 6, which can be compared to the psychophysical results in Figure 5B.

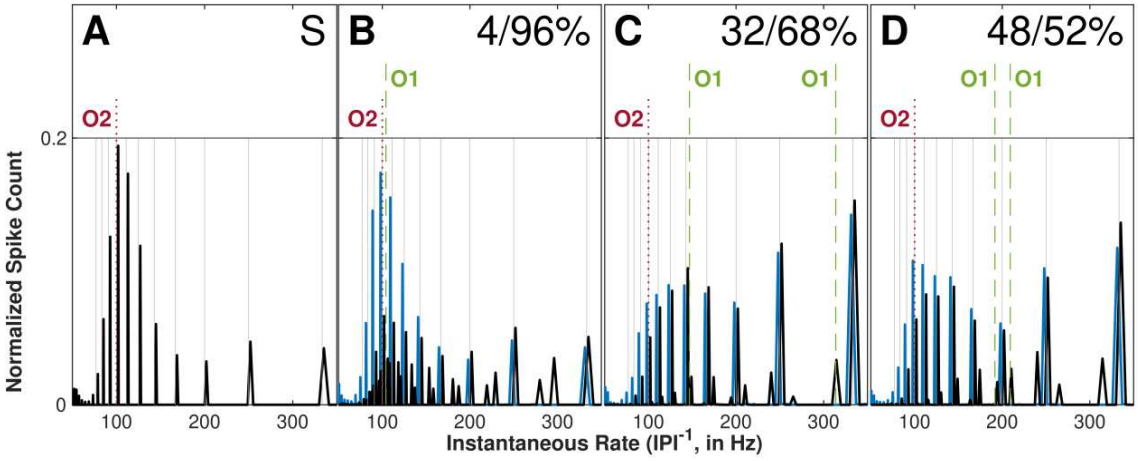


Figure 7

*Normalized spike counts as a function of instantaneous rate for **HR** stimuli. Vertical thin solid grey lines indicate rates corresponding to sub-multiples of the carrier pulse rate. All other aspects as in Figure 6.*

For single-electrode stimulation (panel A), the T0 is still dominantly coded, although with many IPIs corresponding to sub-multiples of the carrier rate (vertical grey lines, see Appendix for details). As summarized in Table 5, both the spike-count-weighted mean rate (M) as an indicator of the effective rate (towards rate-limitation effects) and the spike-count-weighted standard deviation (SD) as a measure of perceptual salience (see Appendix for details) are increased as compared to LR stimulation. Particularly the much larger SD reflects lower HR sensitivity.

For dual-electrode stimulation with zero tonotopic separation (panels B-D, black lines), the T0 is much less dominant, M's are increased by up to 50%, and SD's are considerably larger, all well in line with the lower *d'*s. Beyond considerably spiking in response to carrier sub-multiples, as found for single-electrode stimulation as well, there seems to occur inharmonic distortion, particularly for the 4/96% condition and instantaneous

rates around 300 Hz. Such distortions have been shown to reduce temporal-pitch sensitivity (e.g., Goldsworthy et al., 2022) and may be an additional explanation of reduced dual-electrode HR sensitivity. With respect to IPI-based temporal-pitch theories, it is evident that both O1's and O2's are not prominently coded with HR stimulation. Instead, M and particularly SD seem to be more straightforward parameters to assess pitch coding. Future research may investigate how IPI-based pitch theories can be evaluated with such more complex high-rate stimulation patterns.

For dual-electrode stimulation with small tonotopic separation (panels B-D, blue lines), the model response differs from the zero-separation condition mainly for mTEA of 4/96%. Compared to single-electrode stimulation, SD is increased while M is similar. Still, the model output seems to be well in line with the *d'*s. For the 32/68 and 48/52, small-separation stimulation yielded only slightly lower M's and similar SD's as compared to zero-separation stimulation, which does not reflect the larger *d'*s for small as compared to zero-separation stimulation. Across mTEA, however, inharmonic distortions were largely absent with small-separation stimulation.

To assess why the model did not reflect the differences in perceptual sensitivity between small- and zero-separation, particularly for the 32/68% mTEA, we applied the same changes to the tonotopic separation and the small-separation amplitudes as for LR stimulation above. Details are provided in the Appendix. Conversely to LR stimulation, lowering the per-electrode amplitude rather than the tonotopic separation seems to be the most crucial change to account for small-separation vs. zero-separation stimulation. One explanation for the potential benefit of low per-electrode amplitudes with multi-electrode stimulation may be that with decreasing amplitude more and more pulses around the AM valley stimulate AN neurons around or below their threshold, effectively gating the modulation valleys and, thus, enhancing the AM depth. Stimuli with gating of AM valleys have been shown to yield higher

ITD sensitivity as compared to ungated stimuli (Laback et al., 2011; Monaghan & Seeber, 2016).

Finally, we modeled the neural response to SIPI stimulation to particularly explore potential reasons for the absent mTEA effect in the S_B condition. Details are provided in the Appendix. In summary, the spike count patterns seem to resemble both LR and HR features. LR features are the relatively prominent O1 coding for zero-separation stimulation and the prominent O2 coding for small-separation stimulation. HR features are carrier-sub-multiple stimulation, inharmonic distortions in the zero-separation conditions, and a large effect of per-electrode amplitude but not tonotopic separation for small- vs zero-separation stimulation.

Taken together, our modeling approach seemed to have qualitatively captured many of the perceptual effects. The model output suggests that LR stimulation is most affected by changes in tonotopic separation while HR stimulation is most affected by the per-electrode amplitude. SIPI stimulation seems to share aspects of both LR and HR stimulation, in line with the idea that it provides better temporal sensitivity while still being able to transmit salient envelope cues for speech understanding.

Conclusions

We investigated temporal-pitch discrimination sensitivity with dual-electrode stimulation in five listeners with MED-EL CIs. Per electrode, stimuli were either unmodulated 100-pps pulse trains (LR stimuli), 1000-pps pulse trains with an amplitude modulation of 100 Hz (HR stimuli), or HR stimuli with SIPI pulses inserted at AM peaks (SIPI stimuli). The per-electrode rate of usable temporal information (referred to as F0) was always 100 Hz. We varied the (monaural) temporal electrode asynchrony (mTEA), within one usable stimulus period (carrier or modulator) and the tonotopic separation within the

range from 0 to 9.4 mm.

For small tonotopic separations (below approximately 2.2 mm), we found large mTEA effects, resulting in reduced sensitivity for high asynchrony (towards 50% of the stimulus period). This effect was found across stimulus types. Although high-rate stimuli (HR and SIPI) showed overall worse pitch discrimination, we did not observe a stronger mTEA effect (i.e., an additional detrimental effect on pitch sensitivity) for these stimuli compared to LR stimuli, as might be expected given their higher spatio-temporally pulse density. Increasing the tonotopic separation from 0 to 2.2 mm reduced the detrimental effect of asynchrony. For large separations (approximately 7.1 mm on the basal side and approximately 9.4 mm on the apical side), the mTEA effect was completely absent. All these results are consistent with the hypothesis that the mTEA effect is driven by channel interactions. Interestingly, dual-electrode sensitivity was never better than single-electrode sensitivity for the corresponding pair members, indicating that information was not integrated across electrodes or integration was restricted by the loudness-balancing employed.

Qualitative modeling of zero- and small-separation conditions with a phenomenological AN model well reproduced effects of the mTEA on the effective rate with LR stimulation, consistent with temporal-pitch theories relying on pulse-by-pulse interval rather than autocorrelation analysis, and the reduced prominence of the AM rate (i.e., F0) in the response to HR compared to the LR stimulation. Furthermore, the modeling results suggest that SIPI stimulation shares neural-coding aspects of both LR and HR stimulation that are beneficial for temporal-pitch (and ITD) coding versus speech coding, respectively. In its default configuration, the model could not explain the mTEA for small-separation LR stimulation that had not been observed in previous studies using a ranking task, together highlighting the need for further investigation of the role of the task in temporal-cue coding. We conclude that future stimulation paradigms aiming to minimize channel interactions and

maximizing temporal-pitch sensitivity should stimulate adjacent electrodes with minimal asynchrony in the carrier (in case of LR stimuli) or modulator (in case of HR stimuli). Furthermore, the SIPI approach, after now having been tested on more than one electrode, is still a promising candidate to increase timing-cue sensitivity conveyed with high pulse rates.

Acknowledgments

We thank our CI listeners for their patience and commitment while conducting the experiments. We further thank Michael Mihocic, Tobias Harvey, and Marie-Luise Augsten for assisting with data collection, Michael Smithson for discussions on effect-size confidence intervals, and Marko Takanen for helping with model parameterization and validation.

Author Contributions

ML and BL primarily conceptualized and designed the study. ML collected the data, performed statistical analysis, and conducted the modeling analysis. ML wrote the first draft of the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

Declaration of Conflicting Interests

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Data Availability Statement

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Funding

ML was supported by a DOC Fellowship of the Austrian Academy of Sciences (A-25606) and by the Dr. Anton Oelzelt-Newin Fund of the Austrian Academy of Sciences (A-100315). Additional support was provided to all authors by the Austrian Agency for International Cooperation in Education and Research (MULT 07/2020).

949 **References**

950 Amrhein, V., Greenland, S., & McShane, B. (2019). Scientists rise up against statistical sig-
951 nificance. *Nature*, *567*(7748), 305–307. https://doi.org/10.1038/d41586-019-00857-9
952 Arnoldner, C., Riss, D., Kaider, A., Mair, A., Wagenblast, J., Baumgartner, W.-D., Gstöttner,
953 W., & Hamzavi, J.-S. (2008). The Intensity-Pitch Relation Revisited: Monopolar Ver-
954 sus Bipolar Cochlear Stimulation. *The Laryngoscope*, *118*(9), 1630–1636.
955 https://doi.org/10.1097/mlg.0b013e3181799715
956 Arora, K., Dawson, P., Dowell, R., & Vandali, A. (2009). Electrical stimulation rate effects
957 on speech perception in cochlear implants. *International Journal of Audiology*, *48*,
958 561–567. https://doi.org/10.1080/14992020902858967
959 Bakeman, R. (2005). Recommended effect size statistics for repeated measures designs. *Be-*
960 *havior Research Methods*, *37*(3), 379–384. https://doi.org/10.3758/BF03192707
961 Benjamini, Y., & Hochberg, Y. (1995). Controlling the False Discovery Rate: A Practical and
962 Powerful Approach to Multiple Testing. *Journal of the Royal Statistical Society: Se-*
963 *ries B (Methodological)*, *57*(1), 289–300. https://doi.org/10.1111/j.2517-
964 6161.1995.tb02031.x
965 Black, R. C., Clark, G. M., Tong, Y. C., & Patrick, J. F. (1983). Current Distributions in
966 Cochlear Stimulation. *Annals of the New York Academy of Sciences*, *405*(1), 137–145.
967 https://doi.org/10.1111/j.1749-6632.1983.tb31626.x
968 Boëx, C., De Balthasar, C., Kós, M.-I., & Pelizzone, M. (2003). Electrical field interactions
969 in different cochlear implant systems. *The Journal of the Acoustical Society of Amer-*
970 *ica*, *114*(4), 2049–2057. https://doi.org/10.1121/1.1610451
971 Boulet, J., White, M., & Bruce, I. C. (2016). Temporal considerations for stimulating spiral
972 ganglion neurons with cochlear implants. *Journal of the Association for Research in*
973 *Otolaryngology*, *17*(1), 1–17. https://doi.org/10.1007/s10162-015-0545-5

Buechel, B. D., Hancock, K. E., Chung, Y., & Delgutte, B. (2018). Improved neural coding of ITD with bilateral cochlear implants by introducing short inter-pulse intervals. *Journal of the Association for Research in Otolaryngology*, *19*(6), 681–702. https://doi.org/10.1007/s10162-018-00693-0

Carlyon, R. P., & Deeks, J. M. (2002). Limitations on rate discrimination. *The Journal of the Acoustical Society of America*, *112*(3), 1009–1025. https://doi.org/10.1121/1.1496766

Carlyon, R. P., Deeks, J. M., & McKay, C. M. (2010). The upper limit of temporal pitch for cochlear-implant listeners: Stimulus duration, conditioner pulses, and the number of electrodes stimulated. *The Journal of the Acoustical Society of America*, *127*(3), 1469–1478. https://doi.org/10.1121/1.3291981

Carlyon, R. P., Long, C. J., & Deeks, J. M. (2008). Pulse-rate discrimination by cochlear-implant and normal-hearing listeners with and without binaural cues. *The Journal of the Acoustical Society of America*, *123*(4), 2276–2286. https://doi.org/10.1121/1.2874796

Carlyon, R. P., van Wieringen, A., Long, C. J., Deeks, J. M., & Wouters, J. (2002). Temporal pitch mechanisms in acoustic and electric hearing. *The Journal of the Acoustical Society of America*, *112*(2), 621–633. https://doi.org/10.1121/1.1488660

Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist*, *49*(12), 997–1003. https://doi.org/10.1037/0003-066X.49.12.997

Cosentino, S., Deeks, J. M., & Carlyon, R. P. (2015). Procedural Factors That Affect Psychophysical Measures of Spatial Selectivity in Cochlear Implant Users. *Trends in Hearing*, *19*, 233121651560706. https://doi.org/10.1177/2331216515607067

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorials in Quantitative Methods for Psychology*, *1*(1), 42–45. https://doi.org/10.20982/tqmp.01.1.p042

998 Cutler, A., Dahan, D., & van Donselaar, W. (1997). Prosody in the comprehension of spoken
999 language: A literature review. *Language and Speech*, *40*(2), 141–201.
1000 https://doi.org/10.1177/002383099704000203
1001 Dai, H., & Micheyl, C. (2011). Psychometric functions for pure-tone frequency discrimina-
1002 tion. *The Journal of the Acoustical Society of America*, *130*(1), 263–272.
1003 https://doi.org/10.1121/1.3598448
1004 Darwin, C. J. (2005). Pitch and auditory grouping. In C. J. Plack, R. R. Fay, A. J. Oxenham,
1005 & A. N. Popper (Eds.), *Pitch* (Vol. 24, pp. 278–305). Springer.
1006 https://doi.org/10.1007/0-387-28958-5_8
1007 Darwin, C. J., & Carlyon, R. P. (1995). Auditory Grouping. In *Hearing* (pp. 387–424). Else-
1008 vier. https://doi.org/10.1016/B978-012505626-7/50013-3
1009 de Balthasar, C., Boëx, C., Cosendai, G., Valentini, G., Sigrist, A., & Pelizzone, M. (2003).
1010 Channel interactions with high-rate biphasic electrical stimulation in cochlear implant
1011 subjects. *Hearing Research*, *182*(1–2), 77–87. https://doi.org/10.1016/S0378-
1012 5955(03)00174-6
1013 De Groote, E., Carlyon, R. P., Deeks, J. M., & Macherey, O. (2024). Effects of selective
1014 stimulation of apical electrodes on temporal pitch perception by cochlear implant re-
1015 cipients. *The Journal of the Acoustical Society of America*, *156*(3), 2060–2076.
1016 https://doi.org/10.1121/10.0029023
1017 Dienes, Z., & Mclatchie, N. (2018). Four reasons to prefer Bayesian analyses over signifi-
1018 cance testing. *Psychonomic Bulletin & Review*, *25*(1), 207–218.
1019 https://doi.org/10.3758/s13423-017-1266-z
1020 Dietz, M., Marquardt, T., Salminen, N. H., & McAlpine, D. (2013). Emphasis of spatial cues
1021 in the temporal fine structure during the rising segments of amplitude-modulated

sounds. *Proceedings of the National Academy of Sciences*, *110*(37), 15151–15156. https://doi.org/10.1073/pnas.1309712110

Eddington, D. K., Dobelle, W. H., Brackmann, D. E., Mladejovsky, M. G., & Parkin, J. L. (1978). Auditory prostheses research with multiple channel intracochlear stimulation in man. *The Annals of Otology, Rhinology, and Laryngology*, *87*(6 Pt 2), 1–39.

Egger, K., Majdak, P., & Laback, B. (2016). Channel interaction and current level affect across-electrode integration of interaural time differences in bilateral cochlear-implant listeners. *Journal of the Association for Research in Otolaryngology*, *17*(1), 55–67. https://doi.org/10.1007/s10162-015-0542-8

Favre, E., & Pelizzone, M. (1993). Channel interactions in patients using the Ineraid multi-channel cochlear implant. *Hearing Research*, *66*(2), 150–156. https://doi.org/10.1016/0378-5955(93)90136-O

Fielden, C. A., Kluk, K., & McKay, C. M. (2014). Interpulse interval discrimination within and across channels: Comparison of monopolar and tripolar mode of stimulation. *The Journal of the Acoustical Society of America*, *135*(5), 2913–2922. https://doi.org/10.1121/1.4869687

Francart, T., Lenssen, A., Büchner, A., Lenarz, T., & Wouters, J. (2015). Effect of channel envelope synchrony on interaural time difference sensitivity in bilateral cochlear im-plant listeners. *Ear and Hearing*, *36*(4), e199–e206. https://doi.org/10.1097/AUD.0000000000000152

Friesen, L. M., Shannon, R. V., & Cruz, R. J. (2005). Effects of Stimulation Rate on Speech Recognition with Cochlear Implants. *Audiology and Neurotology*, *10*(3), 169–184. https://doi.org/10.1159/000084027

Fu, Q.-J., Chinchilla, S., & Galvin, J. J. (2004). The Role of Spectral and Temporal Cues in Voice Gender Discrimination by Normal-Hearing Listeners and Cochlear Implant

Users. *Journal of the Association for Research in Otolaryngology*, *5*(3), 253–260. https://doi.org/10.1007/s10162-004-4046-1

Galvin, J. J., Oba, S., Başkent, D., & Fu, Q.-J. (2015). Modulation frequency discrimination with single and multiple channels in cochlear implant users. *Hearing Research*, *324*, 7–18. https://doi.org/10.1016/j.heares.2015.02.007

Gaudrain, E., & Başkent, D. (2018). Discrimination of voice pitch and vocal-tract length in cochlear implant users. *Ear and Hearing*, *39*(2), 226–237. https://doi.org/10.1097/aud.0000000000000480

Gold, J. I., & Ding, L. (2013). How mechanisms of perceptual decision-making affect the psychometric function. *Progress in Neurobiology*, *103*, 98–114. https://doi.org/10.1016/j.pneurobio.2012.05.008

Goldsworthy, R. L., Bissmeyer, S. R. S., & Camarena, A. (2022). Advantages of Pulse Rate Compared to Modulation Frequency for Temporal Pitch Perception in Cochlear Implant Users. *Journal of the Association for Research in Otolaryngology*, *23*(1), 137–150. https://doi.org/10.1007/s10162-021-00828-w

Goldsworthy, R. L., Camarena, A., & Bissmeyer, S. R. S. (2021). Pitch perception is more robust to interference and better resolved when provided by pulse rate than by modulation frequency of cochlear implant stimulation. *Hearing Research*, *409*, 108319. https://doi.org/10.1016/j.heares.2021.108319

Grantham, D. W., Ashmead, D. H., Ricketts, T. A., Haynes, D. S., & Labadie, R. F. (2008). Interaural Time and Level Difference Thresholds for Acoustically Presented Signals in Post-Lingually Deafened Adults Fitted with Bilateral Cochlear Implants Using CIS+ Processing. *Ear and Hearing*, *29*(1), 33–44. https://doi.org/10.1097/AUD.0b013e31815d636f

1071 Griessner, A., Schatzer, R., Steixner, V., Rajan, G. P., Zierhofer, C., & Távora-Vieira, D.
1072 (2021). Temporal Pitch Perception in Cochlear-Implant Users: Channel Independence
1073 in Apical Cochlear Regions. *Trends in Hearing*, *25*, 233121652110206.
1074 https://doi.org/10.1177/23312165211020645
1075 Hancock, K. E., Chung, Y., & Delgutte, B. (2012). Neural ITD coding with bilateral cochlear
1076 implants: Effect of binaurally coherent jitter. *Journal of Neurophysiology*, *108*(3),
1077 714–728. https://doi.org/10.1152/jn.00269.2012
1078 Hancock, K. E., Chung, Y., McKinney, M. F., & Delgutte, B. (2017). Temporal envelope
1079 coding by inferior colliculus neurons with cochlear implant stimulation. *Journal of the*
1080 *Association for Research in Otolaryngology*, *18*(6), 771–788.
1081 https://doi.org/10.1007/s10162-017-0638-4
1082 Hartmann, R., Topp, G., & Klinke, R. (1984). Discharge patterns of cat primary auditory fi-
1083 bers with electrical stimulation of the cochlea. *Hearing Research*, *13*(1), 47–62.
1084 https://doi.org/10.1016/0378-5955(84)90094-7
1085 Heffer, L. F., Sly, D. J., Fallon, J. B., White, M. W., Shepherd, R. K., & O'Leary, S. J.
1086 (2010). Examining the Auditory Nerve Fiber Response to High Rate Cochlear Implant
1087 Stimulation: Chronic Sensorineural Hearing Loss and Facilitation. *Journal of Neuro-*
1088 *physiology*, *104*(6), 3124–3135. https://doi.org/10.1152/jn.00500.2010
1089 Hochmair, I., Nopp, P., Jolly, C., Schmidt, M., ser, H. S., Garnham, C., & Anderson, I.
1090 (2006). MED-EL cochlear implants: State of the art and a glimpse into the future.
1091 *Trends in Amplification*, *10*(4), 201–219. https://doi.org/10.1177/1084713806296720
1092 Horne, C. D. F., Sumner, C. J., & Seeber, B. U. (2016). A Phenomenological Model of the
1093 Electrically Stimulated Auditory Nerve Fiber: Temporal and Biphasic Response Prop-
1094 erties. *Frontiers in Computational Neuroscience*, *10*.
1095 https://doi.org/10.3389/fncom.2016.00008

1096 Houtsma, A. J. M. (1995). Pitch perception. In B. C. J. Moore (Ed.), *Hearing* (pp. 267–295).
1097 Academic Press. https://doi.org/10.1016/B978-012505626-7/50010-8
1098 Hu, H., Ewert, S. D., McAlpine, D., & Dietz, M. (2017). Differences in the temporal course
1099 of interaural time difference sensitivity between acoustic and electric hearing in am-
1100 plitude modulated stimuli. *The Journal of the Acoustical Society of America*, *141*(3),
1101 1862–1873. https://doi.org/10.1121/1.4977014
1102 Huynh, H., & Feldt, L. S. (1970). Conditions under Which Mean Square Ratios in Repeated
1103 Measurements Designs Have Exact F-Distributions. *Journal of the American Statisti-*
1104 *cal Association*, *65*(332), 1582–1589.
1105 https://doi.org/10.1080/01621459.1970.10481187
1106 Ihlefeld, A., Carlyon, R. P., Kan, A., Churchill, T. H., & Litovsky, R. Y. (2015). Limitations
1107 on monaural and binaural temporal processing in bilateral cochlear implant listeners.
1108 *Journal of the Association for Research in Otolaryngology*, *16*(5), 641–652.
1109 https://doi.org/10.1007/s10162-015-0527-7
1110 JASP Team. (2023). *JASP (Version 0.17)*. Retrieved from https://jasp-stats.org/
1111 Kass, R. E., & Raftery, A. E. (1995). Bayes Factors. *Journal of the American Statistical Asso-*
1112 *ciation*, *90*(430), 773–795. https://doi.org/10.1080/01621459.1995.10476572
1113 Kay, M., Elkin, L. A., Higgins, J. J., & Wobbrock, J. O. (2021). *ARTool: Aligned Rank*
1114 *Transform for Nonparametric Factorial ANOVAs* (Version v0.11.0) [Computer soft-
1115 ware]. [object Object]. https://doi.org/10.5281/ZENODO.594511
1116 Keysers, C., Gazzola, V., & Wagenmakers, E.-J. (2020). Using Bayes factor hypothesis test-
1117 ing in neuroscience to establish evidence of absence. *Nature Neuroscience*, *23*(7),
1118 788–799. https://doi.org/10.1038/s41593-020-0660-4

Klein, S. A. (2001). Measuring, estimating, and understanding the psychometric function: A commentary. *Perception & Psychophysics*, *63*(8), 1421–1455. https://doi.org/10.3758/BF03194552

Kong, Y.-Y., Deeks, J. M., Axon, P. R., & Carlyon, R. P. (2009). Limits of temporal pitch in cochlear implants. *The Journal of the Acoustical Society of America*, *125*(3), 1649–1657. https://doi.org/10.1121/1.3068457

Kreft, H. A., Oxenham, A. J., & Nelson, D. A. (2010). Modulation rate discrimination using half-wave rectified and sinusoidally amplitude modulated stimuli in cochlear-implant users. *The Journal of the Acoustical Society of America*, *127*(2), 656–659. https://doi.org/10.1121/1.3282947

Laback, B., & Majdak, P. (2008). Binaural jitter improves interaural time-difference sensitivity of cochlear implantees at high pulse rates. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(2), 814–817. https://doi.org/10.1073/pnas.0709199105

Laback, B., Pok, S.-M., Baumgartner, W.-D., Deutsch, W. A., & Schmid, K. (2004). Sensitivity to interaural level and envelope time differences of two bilateral cochlear implant listeners using clinical sound processors. *Ear and Hearing*, *25*(5), 488–500. https://doi.org/10.1097/01.aud.0000145124.85517.e8

Laback, B., Zimmermann, I., Majdak, P., Baumgartner, W.-D., & Pok, S.-M. (2011). Effects of envelope shape on interaural envelope delay sensitivity in acoustic and electric hearing. *The Journal of the Acoustical Society of America*, *130*(3), 1515–1529. https://doi.org/10.1121/1.3613704

Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00863

1144 Leek, M. R. (2001). Adaptive procedures in psychophysical research. *Perception & Psycho-*
1145 *physics*, *63*(8), 1279–1292. https://doi.org/10.3758/BF03194543
1146 Levitt, H. (1971). Transformed Up-Down Methods in Psychoacoustics. *The Journal of the*
1147 *Acoustical Society of America*, *49*(2), 467–477. https://doi.org/10.1121/1.1912375
1148 Lindenbeck, M. J., Laback, B., Majdak, P., & Srinivasan, S. (2020). Temporal-pitch sensitiv-
1149 ity in electric hearing with amplitude modulation and inserted pulses with short inter-
1150 pulse intervals. *The Journal of the Acoustical Society of America*, *147*(2), 777–793.
1151 https://doi.org/10.1121/10.0000610
1152 Lindenbeck, M. J., Majdak, P., & Laback, B. (2024). Effects of Monaural Temporal Elec-
1153 trode Asynchrony and Channel Interactions in Bilateral and Unilateral Cochlear-Im-
1154 plant Stimulation. *Trends in Hearing*, *28*, 23312165241271340.
1155 https://doi.org/10.1177/23312165241271340
1156 Lindenbeck, M. J., Majdak, P., Srinivasan, S., & Laback, B. (2023). Pitch discrimination in
1157 electric hearing with inconsistent and consistent amplitude-modulation and inter-pulse
1158 rate cues. *The Journal of the Acoustical Society of America*, *153*(6), 3268–3280.
1159 https://doi.org/10.1121/10.0019452
1160 Loizou, P. C., Poroy, O., & Dorman, M. (2000). The effect of parametric variations of coch-
1161 lear implant processors on speech understanding. *The Journal of the Acoustical Soci-*
1162 *ety of America*, *108*(2), 790–802. https://doi.org/10.1121/1.429612
1163 Macherey, O., & Carlyon, R. P. (2010). Temporal pitch percepts elicited by dual-channel
1164 stimulation of a cochlear implant. *The Journal of the Acoustical Society of America*,
1165 *127*(1), 339–349. https://doi.org/10.1121/1.3269042
1166 Majdak, P., Laback, B., & Baumgartner, W.-D. (2006). Effects of interaural time differences
1167 in fine structure and envelope on lateral discrimination in electric hearing. *The*

Journal of the Acoustical Society of America, *120*(4), 2190–2201. https://doi.org/10.1121/1.2258390

McKay, C. M., & Carlyon, R. P. (1999). Dual temporal pitch percepts from acoustic and electric amplitude-modulated pulse trains. *The Journal of the Acoustical Society of America*, *105*(1), 347–357. https://doi.org/10.1121/1.424553

McKay, C. M., & Henshall, K. R. (2009). Amplitude modulation and loudness in cochlear implantees. *Journal of the Association for Research in Otolaryngology*, *11*(1), 101–111. https://doi.org/10.1007/s10162-009-0188-5

McKay, C. M., & McDermott, H. J. (1996). The perception of temporal patterns for electrical stimulation presented at one or two intracochlear sites. *The Journal of the Acoustical Society of America*, *100*(2), 1081–1092. https://doi.org/10.1121/1.416294

McKay, C. M., McDermott, H. J., & Clark, G. M. (1995). Pitch matching of amplitude-modulated current pulse trains by cochlear implantees: The effect of modulation depth. *The Journal of the Acoustical Society of America*, *97*(3), 1777–1785. https://doi.org/10.1121/1.412054

Middlebrooks, J. C. (2004). Effects of cochlear-implant pulse rate and inter-channel timing on channel interactions and thresholds. *The Journal of the Acoustical Society of America*, *116*(1), 452–468. https://doi.org/10.1121/1.1760795

Miller, S. E., Schlauch, R. S., & Watson, P. J. (2010). The effects of fundamental frequency contour manipulations on speech intelligibility in background noise. *The Journal of the Acoustical Society of America*, *128*(1), 435–443. https://doi.org/10.1121/1.3397384

Monaghan, J. J. M., & Seeber, B. U. (2016). A method to enhance the use of interaural time differences for cochlear implants in reverberant environments. *The Journal of the Acoustical Society of America*, *140*(2), 1116–1129. https://doi.org/10.1121/1.4960572

1193 Moore, B. C. J., & Carlyon, R. P. (2005). Perception of pitch by people with cochlear hearing
1194 loss and by cochlear implant users. In C. J. Plack, R. R. Fay, A. J. Oxenham, & A. N.
1195 Popper (Eds.), *Pitch* (Vol. 24, pp. 234–277). Springer. https://doi.org/10.1007/0-387-
1196 28958-5_7
1197 Moore, B. C. J., Glasberg, B. R., & Roberts, B. (1984). Refining the measurement of psycho-
1198 physical tuning curves. *The Journal of the Acoustical Society of America*, *76*(4),
1199 1057–1066. https://doi.org/10.1121/1.391425
1200 Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau
1201 (2005). *Tutorials in Quantitative Methods for Psychology*, *4*(2), 61–64.
1202 https://doi.org/10.20982/tqmp.04.2.p061
1203 Neff, D. L. (1985). Stimulus parameters governing confusion effects in forward masking. *The*
1204 *Journal of the Acoustical Society of America*, *78*(6), 1966–1976.
1205 https://doi.org/10.1121/1.392653
1206 Nelson, D. A., Donaldson, G. S., & Kreft, H. (2008). Forward-masked spatial tuning curves
1207 in cochlear implant users. *The Journal of the Acoustical Society of America*, *123*(3),
1208 1522–1543. https://doi.org/10.1121/1.2836786
1209 Nelson, D. A., Kreft, H. A., Anderson, E. S., & Donaldson, G. S. (2011). Spatial tuning
1210 curves from apical, middle, and basal electrodes in cochlear implant users. *The Jour-*
1211 *nal of the Acoustical Society of America*, *129*(6), 3916–3933.
1212 https://doi.org/10.1121/1.3583503
1213 Oberfeld, D., & Franke, T. (2013). Evaluating the robustness of repeated measures analyses:
1214 The case of small sample sizes and nonnormal data. *Behavior Research Methods*,
1215 *45*(3), 792–812. https://doi.org/10.3758/s13428-012-0281-2

O'Leary, S. J., Black, R. C., & Clark, G. M. (1985). Current distributions in the cat cochlea: A modelling and electrophysiological study. *Hearing Research*, *18*(3), 273–281. https://doi.org/10.1016/0378-5955(85)90044-9

Olejnik, S., & Algina, J. (2003). Generalized Eta and Omega Squared Statistics: Measures of Effect Size for Some Common Research Designs. *Psychological Methods*, *8*(4), 434–447. https://doi.org/10.1037/1082-989X.8.4.434

Oxenham, A. J. (2018). How We Hear: The Perception and Neural Coding of Sound. *Annual Review of Psychology*, *69*(1), 27–50. https://doi.org/10.1146/annurev-psych-122216-011635

Parncutt, R. (1989). *Harmony: A Psychoacoustical Approach* (Vol. 19). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-74831-8

Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of the vowels. *The Journal of the Acoustical Society of America*, *24*(2), 175–184. https://doi.org/10.1121/1.1906875

R Core Team. (2024). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Riss, D., Hamzavi, J.-S., Blineder, M., Honeder, C., Ehrenreich, I., Kaider, A., Baumgartner, W.-D., Gstoettner, W., & Arnoldner, C. (2014). FS4, FS4-p, and FSP: A 4-month crossover study of 3 fine structure sound-coding strategies. *Ear and Hearing*, *35*(6), e272–e281. https://doi.org/10.1097/aud.0000000000000063

Rouder, J. N., Engelhardt, C. R., McCabe, S., & Morey, R. D. (2016). Model comparison in ANOVA. *Psychonomic Bulletin & Review*, *23*(6), 1779–1786. https://doi.org/10.3758/s13423-016-1026-5

1239 Rouder, J. N., Morey, R. D., Speckman, P. L., & Province, J. M. (2012). Default Bayes fac-
1240 tors for ANOVA designs. *Journal of Mathematical Psychology*, *56*(5), 356–374.
1241 https://doi.org/10.1016/j.jmp.2012.08.001
1242 Rouder, J. N., Morey, R. D., Verhagen, J., Swagman, A. R., & Wagenmakers, E.-J. (2017).
1243 Bayesian analysis of factorial designs. *Psychological Methods*, *22*(2), 304–321.
1244 https://doi.org/10.1037/met0000057
1245 Schurzig, D., Timm, M. E., Majdani, O., Lenarz, T., & Rau, T. S. (2021). The Use of Clini-
1246 cally Measurable Cochlear Parameters in Cochlear Implant Surgery as Indicators for
1247 Size, Shape, and Orientation of the Scala Tympani. *Ear & Hearing*, *42*(4), 1034–
1248 1041. https://doi.org/10.1097/AUD.0000000000000998
1249 Sharma, D., Ng, K. K. W., Birznieks, I., & Vickery, R. M. (2022). The burst gap is a periph-
1250 eral temporal code for pitch perception that is shared across audition and touch. *Scien-*
1251 *tific Reports*, *12*(1), 11014. https://doi.org/10.1038/s41598-022-15269-5
1252 Smith, Z. M. (2010). *Improved sensitivity to interaural time differences with the FAST coding*
1253 *strategy*. 11th International Conference on Cochlear Implants and Other Implantable
1254 Auditory Technologies, Stockholm, Sweden.
1255 Smith, Z. M., Delgutte, B., & Oxenham, A. J. (2002). Chimaeric sounds reveal dichotomies
1256 in auditory perception. *Nature*, *416*(6876), 87–90. https://doi.org/10.1038/416087a
1257 Srinivasan, S., Laback, B., & Majdak, P. (2017). Improving interaural time difference sensi-
1258 tivity using short interpulse intervals with vowel-like stimuli in bilateral cochlear im-
1259 plants. *The Journal of the Acoustical Society of America*, *141*, 3973–3974.
1260 https://doi.org/10.1121/1.4989062
1261 Srinivasan, S., Laback, B., Majdak, P., & Arnoldner, C. (2020). Improving interaural time
1262 difference sensitivity using short inter-pulse intervals with amplitude-modulated pulse

trains in bilateral cochlear implants. *Journal of the Association for Research in Oto-laryngology*, *21*, 105--120. https://doi.org/10.1007/s10162-020-00743-6

Srinivasan, S., Laback, B., Majdak, P., & Delgutte, B. (2018). Introducing short interpulse intervals in high-rate pulse trains enhances binaural timing sensitivity in electric hearing. *Journal of the Association for Research in Otolaryngology*, *19*(3), 301–315. https://doi.org/10.1007/s10162-018-0659-7

Takanen, M., & Seeber, B. U. (2022). A Phenomenological Model Reproducing Temporal Response Characteristics of an Electrically Stimulated Auditory Nerve Fiber. *Trends in Hearing*, *26*, 233121652211170. https://doi.org/10.1177/23312165221117079

Takanen, M., Strahl, S., & Schwarz, K. (2024). Insights Into Electrophysiological Metrics of Cochlear Health in Cochlear Implant Users Using a Computational Model. *Journal of the Association for Research in Otolaryngology*, *25*(1), 63–78. https://doi.org/10.1007/s10162-023-00924-z

Tong, Y. C., & Clark, G. M. (1986). Loudness summation, masking, and temporal interaction for sensations produced by electric stimulation of two sites in the human cochlea. *The Journal of the Acoustical Society of America*, *79*(6), 1958–1966. https://doi.org/10.1121/1.393203

Trainor, L. J., Marie, C., Bruce, I. C., & Bidelman, G. M. (2014). Explaining the high voice superiority effect in polyphonic music: Evidence from cortical evoked potentials and peripheral auditory models. *Hearing Research*, *308*, 60–70. https://doi.org/10.1016/j.heares.2013.07.014

van den Bergh, D., van Doorn, J., Marsman, M., Draws, T., Van Kesteren, E.-J., Derks, K., Dablander, F., Gronau, Q. F., Kucharský, Š., Gupta, A. R. K. N., Sarafoglou, A., Voelkel, J. G., Stefan, A., Ly, A., Hinne, M., Matzke, D., & Wagenmakers, E.-J.

(2020). A Tutorial on Conducting and Interpreting a Bayesian ANOVA in JASP. *L'Année Psychologique*, *Vol. 120*(1), 73–96. https://doi.org/10.3917/anpsy1.201.0073

Van Den Bergh, D., Wagenmakers, E.-J., & Aust, F. (2023). Bayesian Repeated-Measures Analysis of Variance: An Updated Methodology Implemented in JASP. *Advances in Methods and Practices in Psychological Science*, *6*(2), 25152459231168024. https://doi.org/10.1177/25152459231168024

Van Den Honert, C., & Stypulkowski, P. H. (1987). Temporal response patterns of single auditory nerve fibers elicited by periodic electrical stimuli. *Hearing Research*, *29*(2–3), 207–222. https://doi.org/10.1016/0378-5955(87)90168-7

Van Eeckhoutte, M., Spirrov, D., & Francart, T. (2018). Comparison between adaptive and adjustment procedures for binaural loudness balancing. *The Journal of the Acoustical Society of America*, *143*(6), 3720–3729. https://doi.org/10.1121/1.5042522

van Hoesel, R. J. M., Böhm, M., Pesch, J., Vandali, A., Battmer, R. D., & Lenarz, T. (2008). Binaural speech unmasking and localization in noise with bilateral cochlear implants using envelope and fine-timing based strategies. *The Journal of the Acoustical Society of America*, *123*(4), 2249–2263. https://doi.org/10.1121/1.2875229

van Wieringen, A., Carlyon, R. P., Long, C. J., & Wouters, J. (2003). Pitch of amplitude-modulated irregular-rate stimuli in acoustic and electric hearing. *The Journal of the Acoustical Society of America*, *114*(3), 1516–1528. https://doi.org/10.1121/1.1577551

Vandali, A. E., Sucher, C., Tsang, D. J., McKay, C. M., Chew, J. W. D., & McDermott, H. J. (2005). Pitch ranking ability of cochlear implant recipients: A comparison of sound-processing strategies. *The Journal of the Acoustical Society of America*, *117*(5), 3126–3138. https://doi.org/10.1121/1.1874632

Vandali, A., Sly, D., Cowan, R., & van Hoesel, R. (2013). Pitch and loudness matching of unmodulated and modulated stimuli in cochlear implantees. *Hearing Research*, *302*, 32–49. https://doi.org/10.1016/j.heares.2013.05.004

Vongphoe, M., & Zeng, F.-G. (2005). Speaker recognition with temporal cues in acoustic and electric hearing. *The Journal of the Acoustical Society of America*, *118*(5), 1055–1061.

Wetzels, R., Grasman, R. P. P. P., & Wagenmakers, E.-J. (2012). A Default Bayesian Hypothesis Test for ANOVA Designs. *The American Statistician*, *66*(2), 104–111. https://doi.org/10.1080/00031305.2012.695956

Williges, B., Jürgens, T., Hu, H., & Dietz, M. (2018). Coherent coding of enhanced interaural cues improves sound localization in noise with bilateral cochlear implants. *Trends in Hearing*, *22*, 1–18. https://doi.org/10.1177/2331216518781746

Wilson, B. S., & Dorman, M. F. (2008). Cochlear implants: A remarkable past and a brilliant future. *Hearing Research*, *242*(1–2), 3–21. https://doi.org/10.1016/j.heares.2008.06.005

Wilson, B. S., Finley, C. C., Lawson, D. T., Wolford, R. D., Eddington, D. K., & Rabinowitz, W. M. (1991). Better speech recognition with cochlear implants. *Nature*, *352*(6332), 236–238. https://doi.org/10.1038/352236a0

Wobbrock, J. O., Findlater, L., Gergle, D., & Higgins, J. J. (2011). The Aligned Rank Transform for Nonparametric Factorial Analyses Using Only ANOVA Procedures. *Proceedings of the ACM Conference on Human Factors in Computing Systems (CHI '11)*, 143–146. https://doi.org/10.1145/1978942.1978963

Yoshimura, H., Moteki, H., Nishio, S., & Usami, S. (2020). Electric-acoustic stimulation with longer electrodes for potential deterioration in low-frequency hearing. *Acta Oto-Laryngologica*, *140*(8), 624–630. https://doi.org/10.1080/00016489.2020.1760351

Zeng, F.-G. (2022). Celebrating the one millionth cochlear implant. *JASA Express Letters*, *2*(7), 077201. https://doi.org/10.1121/10.0012825

Zeng, F.-G., Nie, K., Liu, S., Stickney, G., Del Rio, E., Kong, Y.-Y., & Chen, H. (2004). On the dichotomy in auditory perception between temporal envelope and fine structure cues (L). *The Journal of the Acoustical Society of America*, *116*(3), 1351–1354. https://doi.org/10.1121/1.1777938

Appendix

Parameterization and Validation of the Auditory-Nerve Fiber Simulation

Rationale

We aimed to model the neural response to our single- and dual-electrode stimuli phenomenologically to *qualitatively* assess the validity of the hypothesis of the dominant or effective rate, respectively, which we recently used to explain the mTEA effect with LR stimuli in pitch and ITD-discrimination tasks (Lindenbeck et al., 2024). Furthermore, we aimed to assess whether a similar hypothesis could explain the mTEA effects we found for high-rate (i.e., HR and SIPI) stimuli and how the experimental stimulus parameters, particularly tonotopic separation (i.e., selected electrodes) and amplitudes, might affect our psychophysical results. Importantly, while LR stimuli are physically simple and, hence, any temporal code they might convey could be broadly assessed without simulations, HR and SIPI stimuli are physically more complex and, thus, the neural response may be more distinct from the physical "code". Simulations help to identify rate cues provided and potentially used by the HR and SIPI stimuli. Importantly, we did not aim to model the discrimination task itself since read-out of sensory information and high-level decision processes can have substantial impact on the behavioral responses (e.g., Gold & Ding, 2013) and an evaluation of such confounding effects is beyond the scope of this study. Hence, we focus on qualitative evaluation of basic phenomena observed on average across subjects, making our simulation less dependent on idiosyncratic effects of decoding and decision processes.

Parameterization

We chose the biphasic leaky integrate-and-fire model developed by Horne et al. (2016) with the extension to sequential (i.e., pulse train) stimuli (Takanen & Seeber, 2022). This so-called S-BLIF model was shown to replicate the temporal neural effects of refractoriness, facilitation, accommodation, and spike-rate adaptation as measured in sets of

animal (i.e., mainly cat) data. Furthermore, recently the model parameters were updated so that the model generates eCAP thresholds in the human range (Takanen et al., 2024).

As model input, we used exact replications of the experimental stimuli (see Section "Stimuli and Apparatus" and Figure 1) with an F0 of 100 Hz and amplitudes (in µA) derived from the results of the loudness-balancing procedure. The amplitudes, averaged across listeners, are summarized for single-electrode stimulation in Table 5 and for dual-electrode stimulation in Table 6. For small-separation dual-electrode conditions, amplitudes from apical and basal C2 electrodes were averaged. To better match the model validation data (cf. Takanen & Seeber, 2022), we used a phase duration of 40 µs for the simulations.

To model the distance between electrode and neuron, we followed the approach of Takanen et al. (2024) who calculated, for averaged electrode positions of the FLEX28 array reported by Yoshimura et al. (2020), the distance to the lateral wall based on human temporal-bone CT scans and 3D modeling of the cochleae (Schurzig et al., 2021). Note that, in our cohort, three out of five listeners used the FLEX28 array. Since we simulated only zero- and small-separation dual-electrode conditions which were closely centered around electrode six (i.e., the C1 electrode, cf. Figure 2A), we used a distance of 0.77 mm (average of the 10% and 90% fits reported in Takanen et al., 2024) and employed a 2 dB/mm damping (Black et al., 1983; O'Leary et al., 1985). For each experimental condition, we ran 100 model simulations. These simulations can be thought of either as 100 repeated stimulations of a single auditory-nerve fiber or, because we modeled only a single cochlear location (i.e., the electrode-neuron distance was kept constant and did not follow a distribution), as a single stimulation of 100 nerve fibers with zero spacing between them (i.e., perfect overlap). In other words, all simulations were run from the perspective of (the center of) the C1 electrode. Consequently, to model the average electrode separation of 2.2 mm (cf. Table 1) in small-separation conditions, the amplitude of the C2 electrode pulse train was damped by the same

2 dB/mm as employed above to model the electrode-neuron distance.[13]

In Table 6, it is evident that the loudness-balanced amplitudes are smaller in the small-separation as compared to the zero-separation conditions (columns “Experimental setup”) despite both conditions containing the same number of pulses and stimulating the same cochlear region. We attribute the smaller small-separation amplitudes to the broader excitation of the auditory nerve and, hence, more neural loudness summation (e.g., Tong & Clark, 1986).

Next, before employing the simulation results to discuss our behavioral hypotheses, we validated our approach in two stimulation scenarios.

Validation

For the validation, we focused on single-electrode conditions from the present study and a preceding study. The first scenario was to simulate plausible causes for the differences in single-electrode pitch-discrimination sensitivity as summarized in Figure 3. The second scenario was to simulate the AM-enhancement effect thought to underlie the addition of SIPI pulses on pitch-discrimination sensitivity (Lindenbeck et al., 2020).

Behavioral sensitivity was best for LR stimuli, worst for HR stimuli, and intermediate for SIPI stimuli. Possible reasons for this may be the higher perceived pitch for HR compared to LR stimuli due to the higher neural firing rates at low AM depths (e.g., McKay et al., 1995; Vandali et al., 2013) and higher perceptual salience of the LR rate cue compared to the HR rate cue (e.g., Goldsworthy et al., 2022), often in conjunction with the rate limitation occurring for rates exceeding about 300 pps (e.g., Carlyon et al., 2008; Ihlefeld et al., 2015). We note that rate limitations with MED-EL CIs may occur at higher rates than commonly reported (e.g., De Groote et al., 2024; Kong et al., 2009), but that rate limitations estimated recently by the authors (Lindenbeck et al., 2020, 2023) were found already for rates exceeding about 200 pps. We simulated the model response to the actual single-electrode

stimuli used in the behavioral experiment and with the actual amplitudes derived from the loudness-balancing results. The amplitudes are noted in Table 5.

Table 5

Listener-averaged currents, spike-count-weighted mean (M) instantaneous rate (cf. panels A of Figure 6, Figure 7, and Figure S6) and spike-count-weighted standard deviation (SD) for single-electrode simulation conditions and different stimulus types (cf. Figure 1, left column).

Type	Current (µA)	M	SD
LR	736	100	0
HR	521	129	60
SIPI	500	115	46

To assess the validity of the simulations, Figure 8 shows the peri-stimulus time histograms of the model output in response to LR, HR, and SIPI stimuli, respectively, against a 50-ms snippet of the 300-ms steady state of the physical stimuli. The response to LR stimuli (panel A) shows neural responses to every stimulus pulse with little to no temporal jitter, indicating high phase locking and negligible neural interplay between successive pulses, which is in line with neurophysiological data from anesthetized cats (Hancock et al., 2017) as well as unanesthetized rabbits (Buechel et al., 2018), and the general notion that the time scales of the temporal effects of refractoriness, facilitation, and accommodation are smaller than the 10-ms IPI in the LR stimuli (e.g., Boulet et al., 2016; Heffer et al., 2010; Takanen & Seeber, 2022).

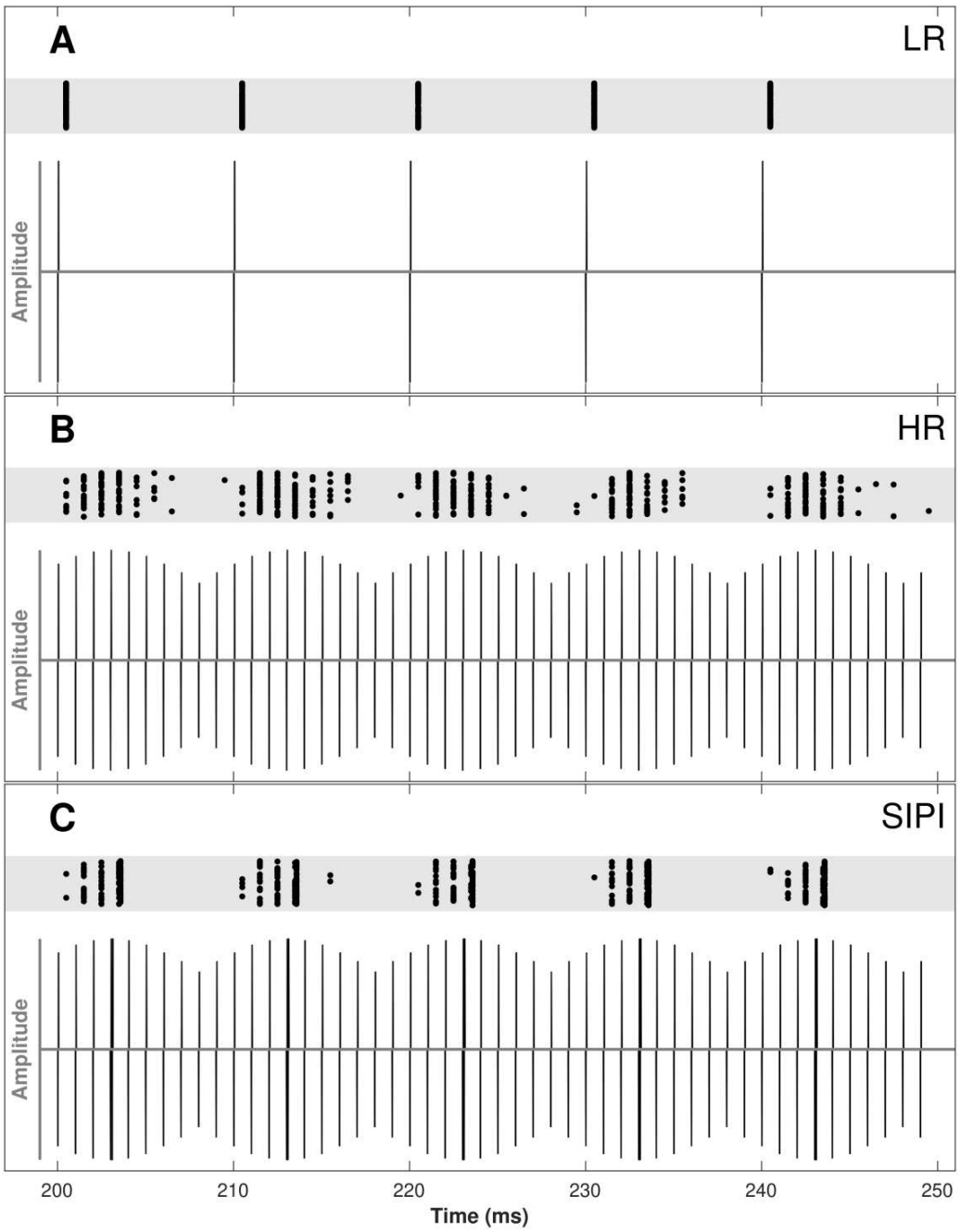


Figure 8

Spikes as a function of time and in response to different single-electrode stimuli (cf. Figure 1, left column) in different panels. For better readability, only a 50-ms snippet of the 300-ms steady state is shown. To better visualize the simulation repetitions, the spikes are randomly jittered on the ordinate within the grey ribbons.

Table 6

Listener-averaged currents for dual-electrode simulation conditions (cf. Table 4): different stimulus types and mTEAs (cf. Figure 1, right column). Amplitude changes Δ as linear differences to either single-electrode (S, cf. Table 5) or dual-electrode zero-separation conditions (Z).

Type	mTEA	**Zero separation (Z)**		**Small separation**							
		Experimental setup		*Experimental setup*		*0.75 mm Separation*		*Amps increased to Z level*		*Small-Zero Amp. Difference doubled*	
		Current (µA)	Δ *re* S	Current (µA)	Δ *re* Z	Current (µA)	Δ *re* Z	Current (µA)	Δ *re* Z	Current (µA)	Δ *re* Z
LR	4/96%	692	-44	643	-49	643	-49	692	0	594	-98
	32/68%	668	-68	636	-32	636	-32	668	0	604	-64
	48/52%	647	-89	628	-19	628	-19	647	0	609	-38
HR	4/96%	501	-20	464	-37	464	-37	501	0	427	-74
	32/68%	485	-36	457	-28	457	-28	485	0	429	-56
	48/52%	495	-26	461	-34	461	-34	495	0	427	-68
SIPI	4/96%	475	-25	438	-37	438	-37	475	0	401	-74
	32/68%	474	-26	433	-41	433	-41	474	0	392	-82
	48/52%	471	-29	435	-36	435	-36	471	0	399	-72

The response to HR stimuli (panel B) shows rough coding of the AM in that most spikes tended to be elicited around AM peaks with slightly more spikes being elicited towards the onset as compared to the offset. Hence, it is evident that the carrier rate crucially affects the temporal code with such stimuli. The coding of the AM peak and the representation of the carrier rate in the neural code to HR stimuli is well in line with neurophysiological data (Hancock et al., 2017) and recent psychophysical data on ITD sensitivity (Hu et al., 2017; Srinivasan et al., 2020).

The response to SIPI stimuli (panel B) shows many parallels to HR coding, with two notable exceptions. First, as shown before (Buechel et al., 2018; Hancock et al., 2012), SIPIs elicit precise spiking at the AM peaks where they are placed. In particular, the simulation even models the fact that sometimes the carrier pulse directly preceding the SIPI pulse elicits the spike, rendering the SIPI pulse ineffective due to refractoriness in the fiber. However, most of the spikes occur in response to the SIPI pulse and, hence, likely due to neural facilitation (or temporal summation). Second, the simulation predicts that neural activity after a SIPI-based spike and, hence, at the AM offset, is largely suppressed, most likely due to many fibers simultaneously being in a SIPI-spike-induced state of refractoriness (not the generally high degree of across-fiber synchrony in electric stimulation; Hartmann et al., 1984). Consequently, the model suggests that the hypothesized SIPI-based AM enhancement effect (Buechel et al., 2018; Hancock et al., 2012; Lindenbeck et al., 2020; Srinivasan et al., 2018) is not simply an enhancement of the AM peak, but a peak enhancement combined with an offset suppression and little changes at the onset. Since human temporal sensitivity is best at the AM onset or peak (e.g., Dietz et al., 2013; Hu et al., 2017; Srinivasan et al., 2020), these SIPI properties may be highly beneficial for temporal coding in electric hearing.

Single-electrode sensitivity

To visualize the periodicities coded by the LR, HR, and SIPI stimuli, Figure 9 shows the normalized spike count as a function of the instantaneous rate (i.e., the inverse of the IPI). We chose an IPI-based analysis instead of an autocorrelation-based one for two reasons. First, IPI-based pitch theories could explain pitch perception with LR stimuli while autocorrelation-based theories could not (Carlyon et al., 2002; McKay & McDermott, 1996; van Wieringen et al., 2003). Second, the (dominant) IPI was shown to be a code for temporal pitch perception that is shared across audition and touch (Sharma et al., 2022). For the spike count distributions underlying Figure 9, we calculated the spike-count-weighted mean

instantaneous rate (M) as indicator of the effective rate towards assessing potential rate-limitation effects and the spike-count-weighted standard deviation (SD) as an indicator of the rate-cue salience (small values indicating higher salience). Table 5 summarizes these parameters.

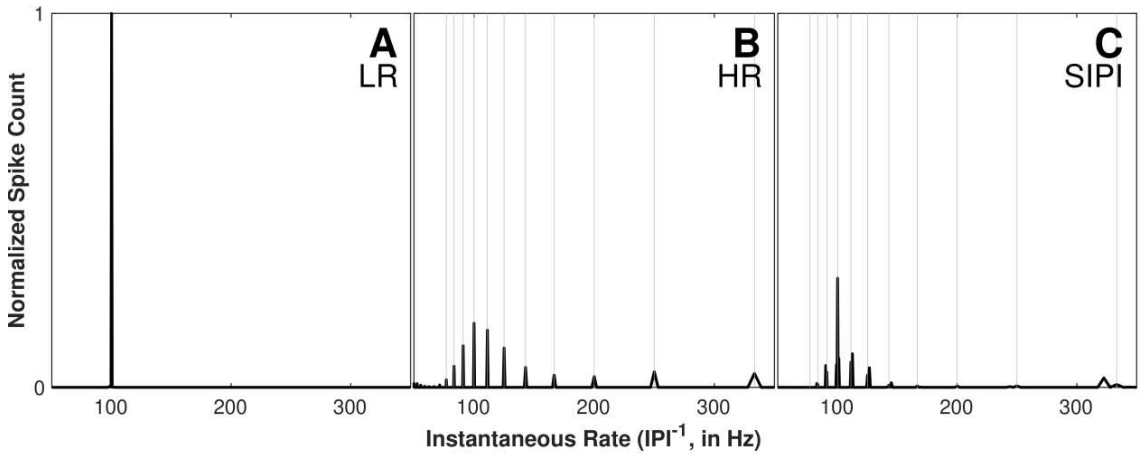


Figure 9

*Spike counts as a function of instantaneous rate for single-electrode stimulation (cf. Figure 1, left column). The data are the same as in Figure 8. To preserve the main effect of stimulus type (cf. Figure 3), counts are normalized to the maximum across conditions. In panels **B** and **C**, vertical thin solid grey lines indicate rates corresponding to sub-multiples of the carrier pulse rate. Table 5 summarizes the weighted average instantaneous rates.*

In Figure 9, for LR stimulation (panel A), the effective rate precisely matches the stimulus pulse rate. The SD of zero quantifies high salience, although probably overestimating it (i.e., the SD indicates a spiking probability of 1 which, despite the strong phase locking in electric hearing, will not be reached under real-life conditions). For HR stimulation (panel B), the IPI distribution is much broader with many spikes occurring in response to carrier pulses (vertical gray lines). Still, the most common IPI corresponds to the F0. Yet, M is increased by 29% and the SD is the largest among the three stimulus types, indicating the lowest salience. For SIPI stimulation (panel C), the IPI distribution is

considerably narrower than for HR stimulation and the F0 is more prominently coded. Both M and SD are intermediate to LR and HR stimulation. Overall, both HR and SIPI stimuli yield several times more spikes than LR stimuli.

In general, we consider the model output to be–apart from possible overestimation of LR salience–well in line with literature on the effective rate with HR as compared to LR stimuli (e.g., Vandali et al., 2013) and differences in cue salience (i.e., SD in model response) to well correspond to the differences in behavioral single-electrode sensitivity (Figure 3).

SIPI-based AM enhancement

We already showed that the simulations predicted precise SIPI-induced, facilitation-driven spiking at AM peaks in conjunction with pronounced refractory effects at AM offsets (cf. Figure 8C). As a final validation step, we aimed to qualitatively replicate the results by Lindenbeck et al. (2020) who showed that HR-based pitch-discrimination sensitivity for AM depths of 0.5 and more is similar to SIPI-based sensitivity (regardless of the AM depth).[14] To this end, we simulated neural responses to single-electrode HR stimuli with AM depths of 0.3, 0.5, 0.7, and 0.9. Since loudness-adjustment results in constant peak amplitude across AM depths (Lindenbeck et al., 2020; McKay & Henshall, 2009), the same peak amplitude as for the HR stimulus with a depth of 0.3 was used for all other depths. Table 7 summarizes the used amplitudes as well as M and SD parameters. Figure 10 furthermore visualizes the spike count as a function of instantaneous rate. With increasing AM depth, the HR IPI distributions become narrower, that is, the SD decreases. However, F0 remains better coded with SIPI stimuli regardless of HR AM depth, that is, M is closer to the F0 for all SIPI conditions as compared to the HR condition. Notably, the M and SD parameters are remarkably similar for the SIPI/0.3-depth stimulus and the HR/0.5-depth stimulus. For HR, SD is larger at the depth of 0.3 and smaller at depths of 0.7 and 0.9, respectively. These modeling results parallel the behavioral results in that both M and SD are similar for SIPI and HR for an HR AM depth of

0.5. Further improvements in M and SD parameters for HR stimuli with larger depths may well be counteracted in the behavioral data by non-stimulus-driven effects such as saturation of *d'* scores, task mental load, etc. Such effects would occur for all stimulus types, in particular LR stimuli, and could, hence, explain the similar sensitivity for LR, SIPI, and HR stimuli for AM depths of 0.5 and larger in Lindenbeck et al. (2020).

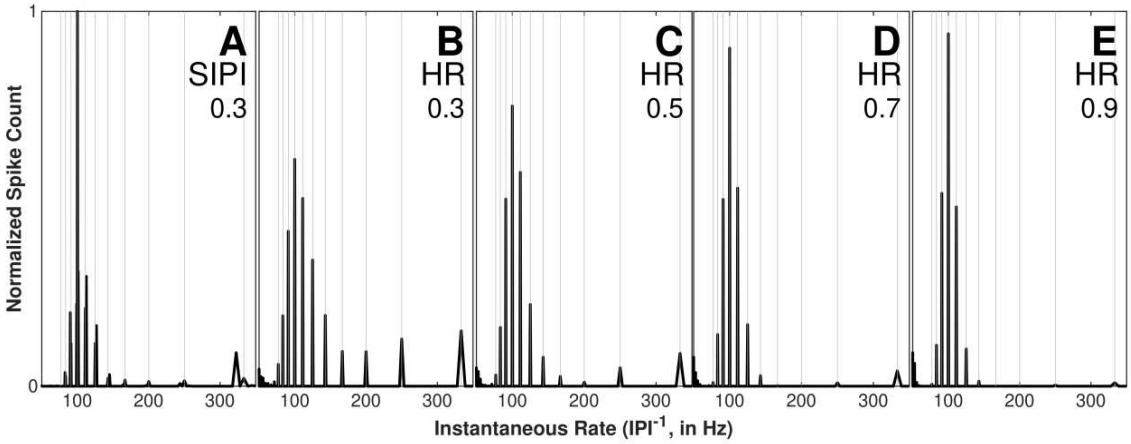


Figure 10

*Spike counts as a function of instantaneous rate for single-electrode SIPI and HR conditions. Panels **A** and **B**: Conditions with an AM depth of 0.3 (cf. Figure 9, panels C and B, respectively) as used in the psychophysical experiment. Panels **C** to **E**: HR condition with increased AM depths of 0.5, 0.7, and 0.9, respectively. Table 7 summarizes the weighted average instantaneous rates. All other aspects as in Figure 9.*

Overall, we consider the validation results to support our model parameterization for further investigations of neural responses to dual-electrode stimulation.

Varying tonotopic separation and small-separation amplitudes

To better understand inconsistencies between the model output and psychophysical results pattern (see Discussion), particularly in small-separation conditions, we repeated the simulations with one modified stimulus parameter. First, to test the effect of the tonotopic separation detached from loudness summation (cf. Tong & Clark, 1986), we included a

condition with 0.75 mm separation but the original 2.2-mm-separation amplitudes (column "0.75 separation"). Second, to simulate loudness summation detached from tonotopic separation, we modified the small-separation amplitudes while keeping the tonotopic separation at 2.2 mm., In particular, we either erased the summation effect by using the zero-separation amplitudes in simulated small-separation conditions (column "Amps increased to Z level") or doubled the summation effect by doubling the amplitude differences between small- and zero-separation conditions (column "Small-Zero Amp. Difference doubled"). Table 6 summarizes the amplitudes for all additional conditions. Figures S4 and S7 summarize the results that are discussed below.

Table 7

Listener-averaged currents (cf. Table 5) and spike-count-weighted mean (M) instantaneous rate and standard deviation (SD) for the SIPI and HR stimuli in Figure 10.

Type	AM Depth	Current (μA)	M	SD
SIPI	0.3	500	115	46
HR	0.3	521	130	62
	0.5		115	49
	0.7		107	34
	0.9		102	19

LR stimulation

Figure S4 shows the LR results. Reducing the simulated tonotopic separation from 2.2 mm to 0.75 mm better modeled the small-separation sensitivity in that fewer IPIs corresponded to the T0, suggesting reduced perceptual salience. Changes in amplitude, however, had little effect on the model output. Taken together, these results suggest that loudness summation has little effect on LR sensitivity while the tonotopic separation directly

affects the rate code and, hence, any rate-associated effects.

HR stimulation

Figure S5 shows the HR results. Since HR-based rate cues are less straightforward to derive from the physical stimulus, we introduced the parameters M and SD to better characterize potential cues. Table 4 summarizes M and SD for the additional simulations. Conversely to LR stimulation, reducing the tonotopic separation had little effect on M and SD. Instead, doubling the small-separation amplitude reduction (i.e., the simulated summation effect) better approximated the pattern of perceptual sensitivity than the default model configuration.

SIPI stimulation

Although we did not analyze psychophysical SIPI data in depth for small tonotopic separations due to the absent mTEA effect in the S_B condition (see Section "Single-electrode vs dual-electrode stimulation for small tonotopic separations"), we ran the simulations both for the default configuration (see Figure S6) and the modified configurations (see Figure S7). The rationale for these additional simulations was to find potential explanations for the absent mTEA effect in the S_B condition.

As shown in Figure S6, the SIPI spike count patterns seem to resemble both LR and HR features. LR features are the relatively prominent O1 coding for zero-separation stimulation (panels B-C, black lines) and the prominent O2 coding for small-separation stimulation. HR features are carrier-sub-multiple stimulation (vertical grey lines) and some inharmonic distortion, particularly in the zero-separation conditions. The prominent O2 coding may reflect absence of channel interactions and, hence, provide an explanation for the absent mTEA effect in the S_B condition, however, this near-full release from channel interactions may also be overestimated in the model output as was the case for LR stimuli.

As shown in Figure S7, when varying the modeled tonotopic separation or loudness

summation, respectively, SIPI stimulation seems to be more closely related to HR stimulation, being only affected by changes in amplitudes but not by changes in tonotopic separation.

Taken together, SIPI simulation seems to combine beneficial features of LR and HR stimulation in dual-electrode configurations, in line with the idea that it might be a suitable approach to better combine temporal and speech coding (cf., e.g., Lindenbeck et al., 2020, 2023; Srinivasan et al., 2018, 2020).

1592 **Footnotes**

[1] Portions of this work were presented at the 47th Annual Meeting of the German Acoustical Society, August 2021, Vienna, Austria, the Conference on Implantable Auditory Prostheses, July 2021, Lake Tahoe, CA, and the 19th International Symposium on Hearing, June 2022, Lyon, France.

[2] Electronic mail: martin.lindenbeck@oeaw.ac.at

[3] Electronic mail: bernhard.laback@oeaw.ac.at

[4] The exact delays deviated negligibly due to technical reasons. The exact delays were 3.96, 31.87, 47.82, 52.18, 68.13, and 96.04%.

[5] As masker and target had the same pulse rate, the target might have been perceived as a continuation of the masker in case both pulse trains were presented to the same electrode. This might have increased the sharpness of the tuning curves (Cosentino et al., 2015; Moore et al., 1984; Neff, 1985). Yet, the selection of C2 electrodes was not based on sharpness.

[6] Note that the *F*-statistic and *p*-value reported here are from an rmANOVA on aligned transformed ranks (ART-rmANOVA; Wobbrock et al., 2011), which is a non-parametric alternative for factorial NHST rmANOVAs with violated assumptions (here, sphericity). In the actual rmANOVA, the *F*-test for the Type effect was marginally insignificant ($p = .056$) due to the Hyunh-Feldt correction (uncorrected $p = .031$). However, the BF_{incl} indicated evidence for a Type effect, making Bayesian and NHST inference inconsistent. Since Bayesian rmANOVAs do not require sphericity (e.g., Rouder et al., 2012), this led us to double-check the NHST result with the ART-ANOVA and report its result consistent with the BF_{incl}. The ART-ANOVA was performed in R 4.3.3 (R Core Team, 2024) using the ARTool package (Kay et al., 2021).

[7] Pairwise comparisons were performed on the aligned transformed ranks.

[8] *F*-statistic and *p*-value are from an ART-rmANOVA, see footnote 6.

[9] Note that the BF is a *continuous* measure of evidence.

[10] *F*-statistic and *p*-value are from an ART-rmANOVA, see footnote 6.

[11] Note that while the Huynh-Feldt correction remedies violations of ANOVA assumptions, the Bonferroni-Holm correction remedies accumulation of Type I errors. Hence, the latter correction does not apply to effect-size confidence intervals.

[12] Post-hoc tests are not richly developed for Bayesian rmANOVAs. Thus, and because we were only exploring differences (i.e., an H1) and not similarity (i.e., an H0), only NHST results are reported.

[13] Because all simulations were run from the perspective of the C1 electrode without any variation in cochlear location of simulated fibers, we always used only the C1 amplitudes from the small-separation loudness-balancing results and applied the damping to these amplitudes to model the C2 amplitudes instead of using the actual C2 amplitudes from the loudness-balancing results. We note that the differences between the loudness-balanced C1 and C2 amplitudes were negligible.

[14] Note that there were several differences in the stimuli between the current study and Lindenbeck et al. (2020): The F0 was 125 Hz instead of 100 Hz, the carrier rate was 2000 pps instead of 1000 pps, electrode eight was used instead of six, the loudness-balanced amplitudes differed, and a different CI listener cohort was tested.

Supplementary Material

Manuscript title:	Effects of Monaural Temporal Electrode Asynchrony on Temporal Pitch Discrimination with Cochlear Implants
Authors:	Martin J. Lindenbeck[a], Piotr Majdak, and Bernhard Laback[b]
Date:	March 6, 2025

[a] martin.lindenbeck@oeaw.ac.at

[b] bernhard.laback@oeaw.ac.at

Tables

Table S1. Inconclusive results of the omnibus rmANOVA conducted to jointly assess the effects of mTEA, CCC, and Type in dual-electrode stimulation. BF_{incl}…Bayes factor; F…test statistic; df_{mTEA}…Huynh-Feldt-adjusted nominator degrees of freedom; df_{error}…Huynh-Feldt-adjusted denominator degrees of freedom; p…Huynh-Feldt-adjusted significance level; η_G^2…effect size estimate [90% confidence interval].

Effect	**BF_{incl}**	***F***	***df_{effect}***	***df_{error}***	***p***	**η_G^2 [5%, 95%]**
Type	2.12	4.30	1.5	5.8	.078	.24 [.01, .44]
CCC × Type	0.60	1.32	8.0	32.0	.268	.05 [.00, .06]
mTEA × CCC × Type	1.69	1.35	27.2	108.7	.142	.03 [.00, .02]

Table S2. Inconclusive results of the rmANOVA conducted to assess the effects of mTEA and Type in dual-electrode stimulation for large-separation CCCs L_A and L_B. All other aspects as in Table S1.

Effect	**BF_{incl}**	***F***	***df_{effect}***	***df_{error}***	***p***	**η_G^2 [5%, 95%]**
CCC	0.80	1.47	1.0	4.0	.293	.02 [.00, .11]
Type	2.16	4.78	1.5	6.1	.062	.29 [.02, .49]
CCC × Type	0.87	1.21	2.0	8.0	.348	.03 [.00, .09]
mTEA × CCC × Type	0.56	0.68	3.2	12.9	.590	.01 [.00, .02]

Table S3. Results of the simple-effects rmANOVAs conducted to assess the mTEA effect separately per CCC. Zero tonotopic separation: CCC Z. Small tonotopic separation: CCCs S_A and S_B. Large tonotopic separation: CCCs L_A and L_B. All other aspects as in Table S1.

CCC	BF_{incl}	F	df_{mTEA}	df_{error}	p	η_G^2 [5%, 95%]
Z	33.81	6.60	1.7	6.7	.029	.26 [.05, .42]
S_A	50.25	6.70	2.5	9.9	.011	.29 [.08, .43]
S_B	3.65	5.91	5.0	20.0	.002	.08 [.03, .11]
L_A	0.12	0.57	5.0	20.0	.726	.01 [.00, .01]
L_B	0.09	0.52	5.0	20.0	.760	.00 [.00, .01]

Table S4. Inconclusive results of the rmANOVA conducted to assess the effects of mTEA and Type in dual-electrode stimulation for zero- and small-separation CCCs Z, S_A, and S_B. All other aspects as in Table S1.

Effect	BF_{incl}	F	df_{effect}	df_{error}	p	η_G^2 [5%, 95%]
Type	1.70	3.60	1.3	5.0	.114	.20 [.00, .42]
mTEA × Type	0.98	2.43	6.1	24.4	.055	.03 [.00, .04]
CCC × Type	0.54	1.01	4.0	16.0	.432	.04 [.00, .08]
mTEA × CCC × Type	0.62	1.19	13.7	54.9	.310	.03 [.00, .03]

Table S5. Inconclusive results of the rmANOVA conducted to assess to the differences between single- and dual-electrode stimulation, zero and small tonotopic separation, certain mTEAs and LR/HR stimuli. All other aspects as in Table S1.

Effect	**BF_{incl}**	***F***	**df_{effect}**	**df_{error}**	***p***	**η_G^2 [5%, 95%]**
Separation × Type	0.65	1.24	1.0	4.0	.328	.03 [.00, .14]
Condition × Type	0.42	1.23	3.0	12.0	.343	.03 [.00, .08]

Table S6. Results of the rmANOVA conducted to assess to the differences between single- and dual-electrode stimulation, zero and small tonotopic separation, and certain mTEAs for **LR** stimuli only. All other aspects as in Table S1.

Effect	**BF_{incl}**	***F***	**df_{effect}**	**df_{error}**	***p***	**η_G^2 [5%, 95%]**
Condition	42.13	10.65	2.3	9.3	.003	.51 [.21, 63]
Separation	.068	1.40	1.0	4.0	.302	.05 [.00, .20]
Condition × Separation	8.31	6.88	1.8	7.3	.022	.12 [.03, .21]

Table S7. Results of the rmANOVA conducted to assess to the differences between single- and dual-electrode stimulation, zero and small tonotopic separation, and certain mTEAs for **HR** stimuli only. All other aspects as in Table S1.

Effect	**BF_{incl}**	***F***	**df_{effect}**	**df_{error}**	***p***	**η_G^2 [5%, 95%]**
Condition (ART ANOVA)	3.11	4.90	3.0	12.0	.019	.26 [.02, .45]
Separation	2.53	13.09	1.0	4.0	.022	.21 [.06, .35]
Condition × Separation	10.36	6.88	2.2	8.8	.015	.12 [.03, .20]

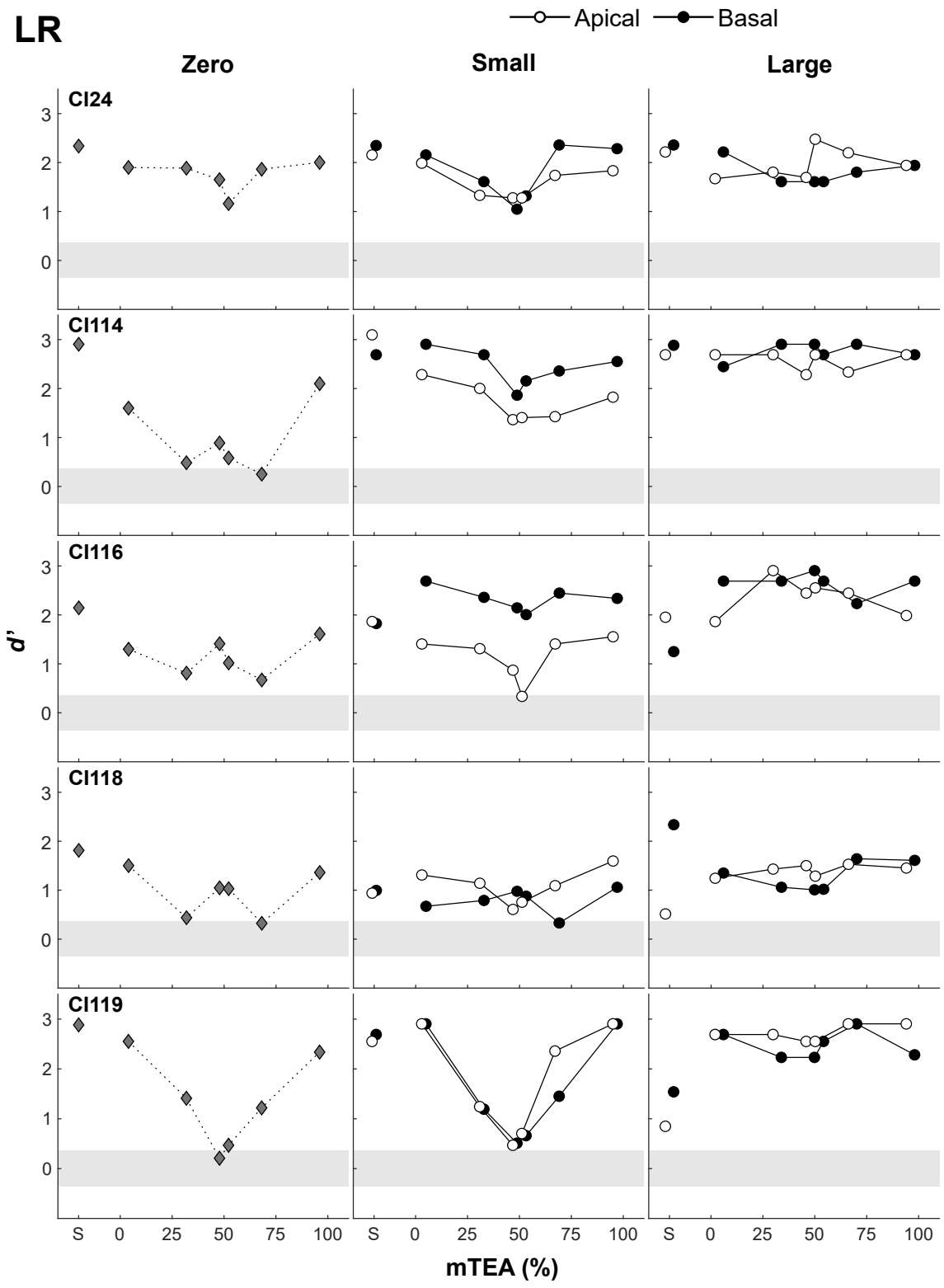


Figure S1

*F0-based pitch-discrimination sensitivity d' for **LR stimuli** as a function of mTEA in separate rows per CI listener. All other aspects as in Figure 3.*

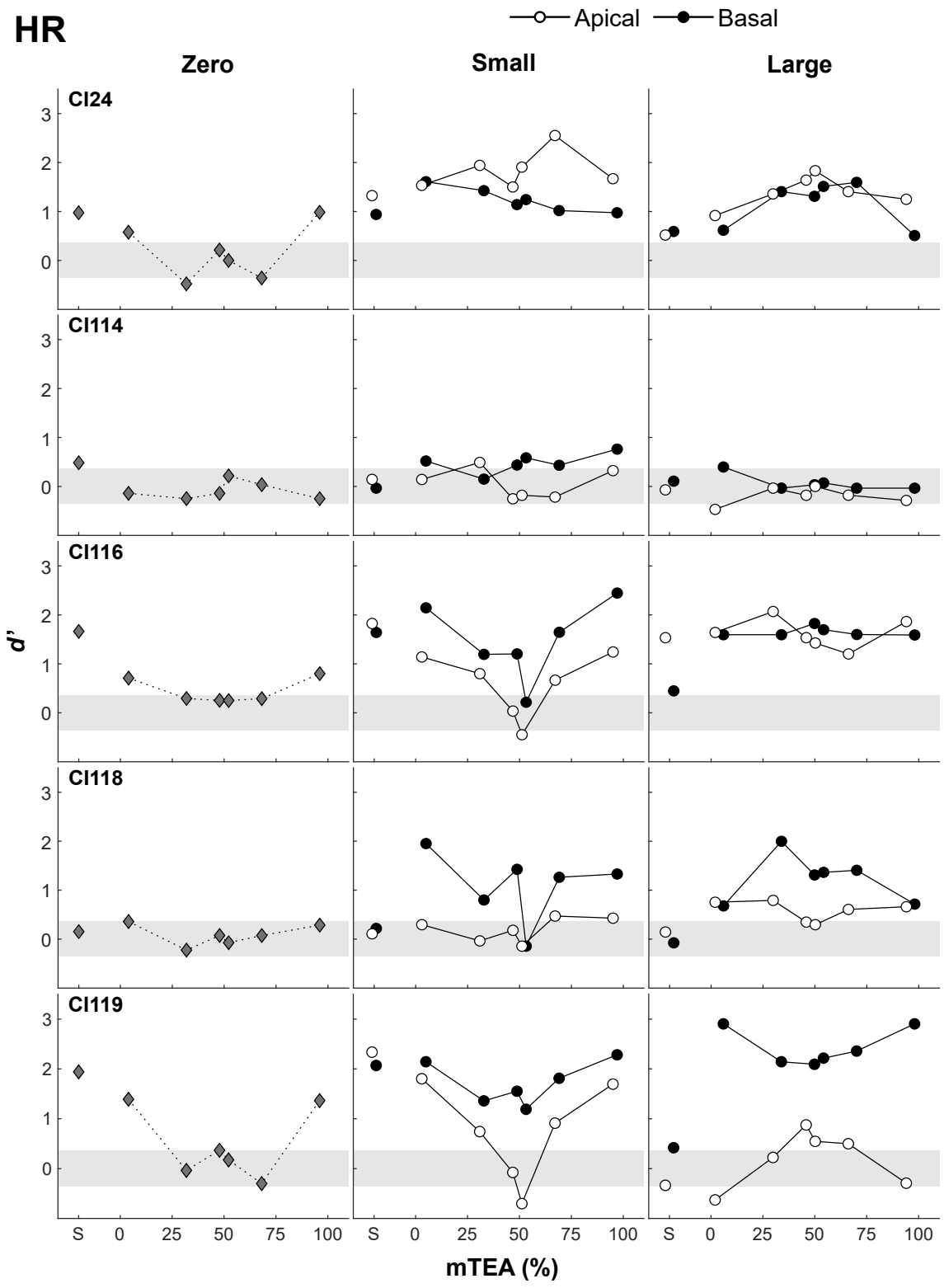


Figure S2

*F0-based pitch-discrimination sensitivity d′ for **HR stimuli** as a function of mTEA in separate rows per CI listener. All other aspects as in Figure 3.*

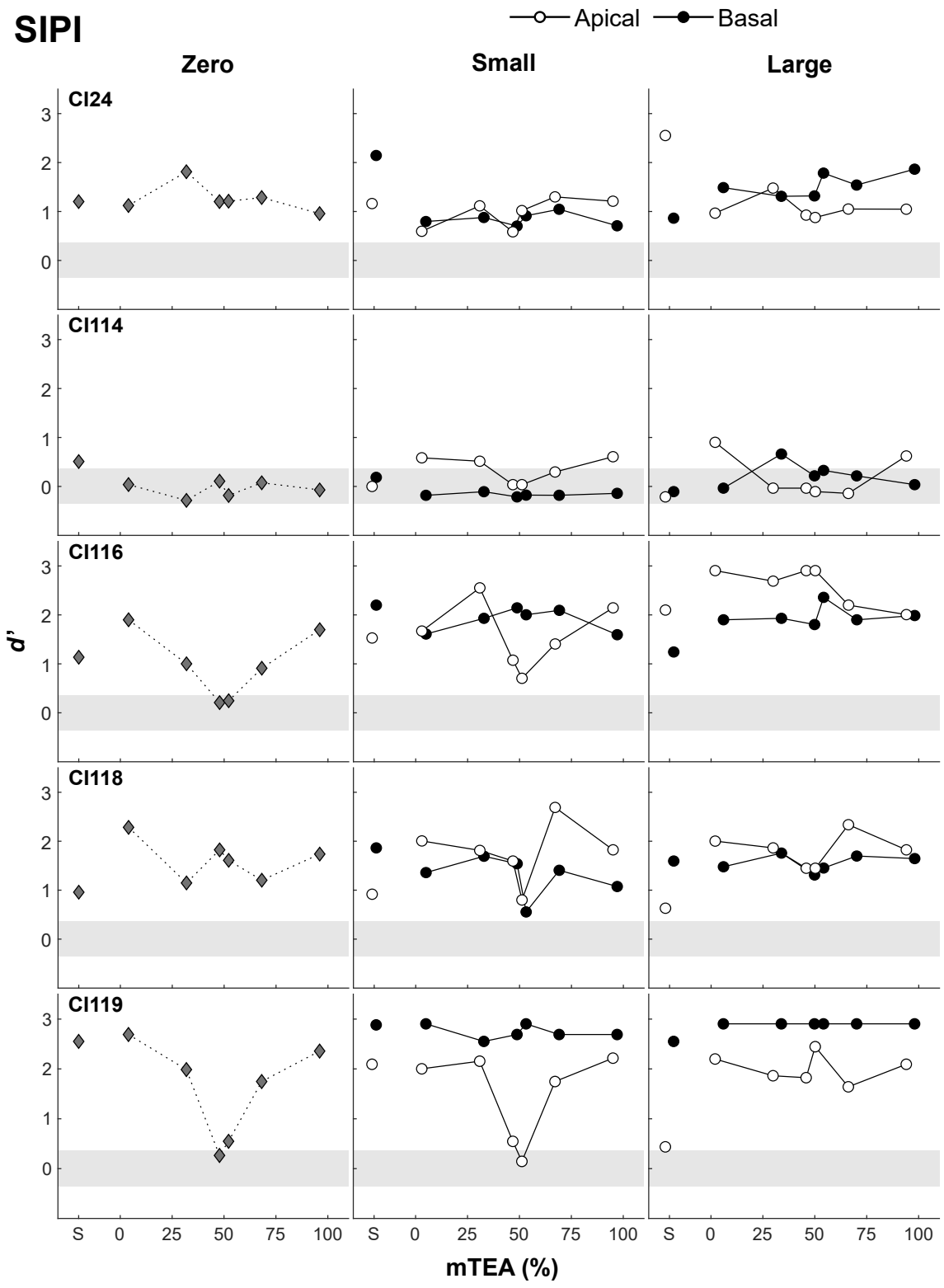
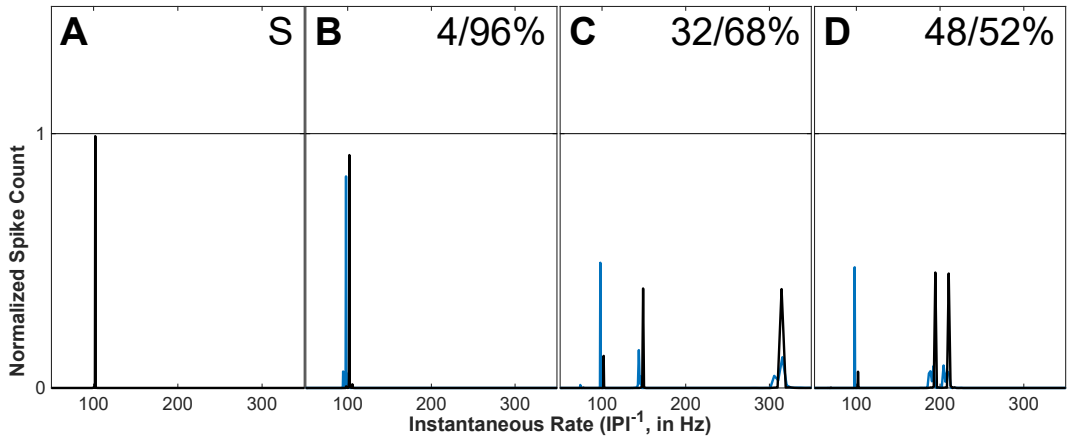


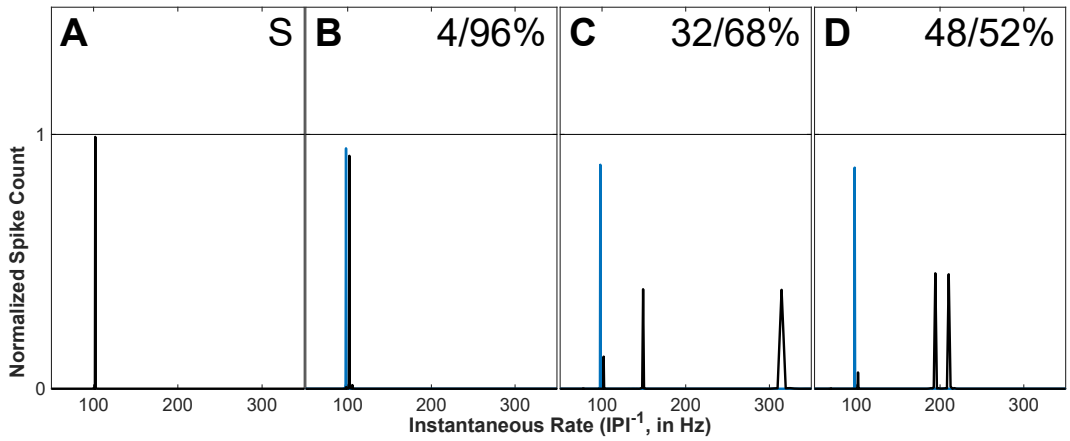
Figure S3

*F0-based pitch-discrimination sensitivity d' for **SIPI stimuli** as a function of mTEA in separate rows per CI listener. All other aspects as in Figure 3.*

Tonotopic Separation: 0.75 mm



S_A/S_B Amplitudes @ Z Level



S_A/S_B Amplitude Reduction re Z *doubled*

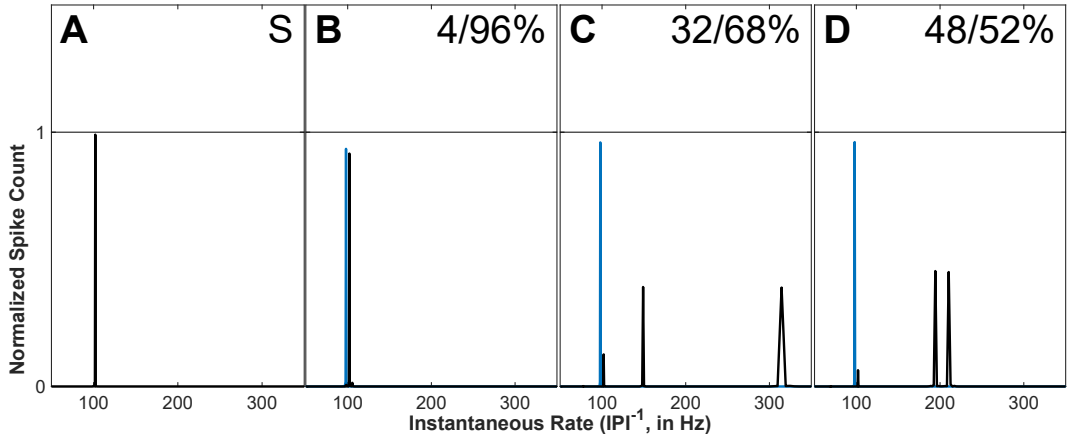


Figure S4

*Spike counts as a function of instantaneous rate (i.e., IPI^{-1}, in Hz) for **LR** with three variations of model parameters as compared to the psychophysical experiment setup (cf. Table 4). All other as-pects as in Figure 7.*

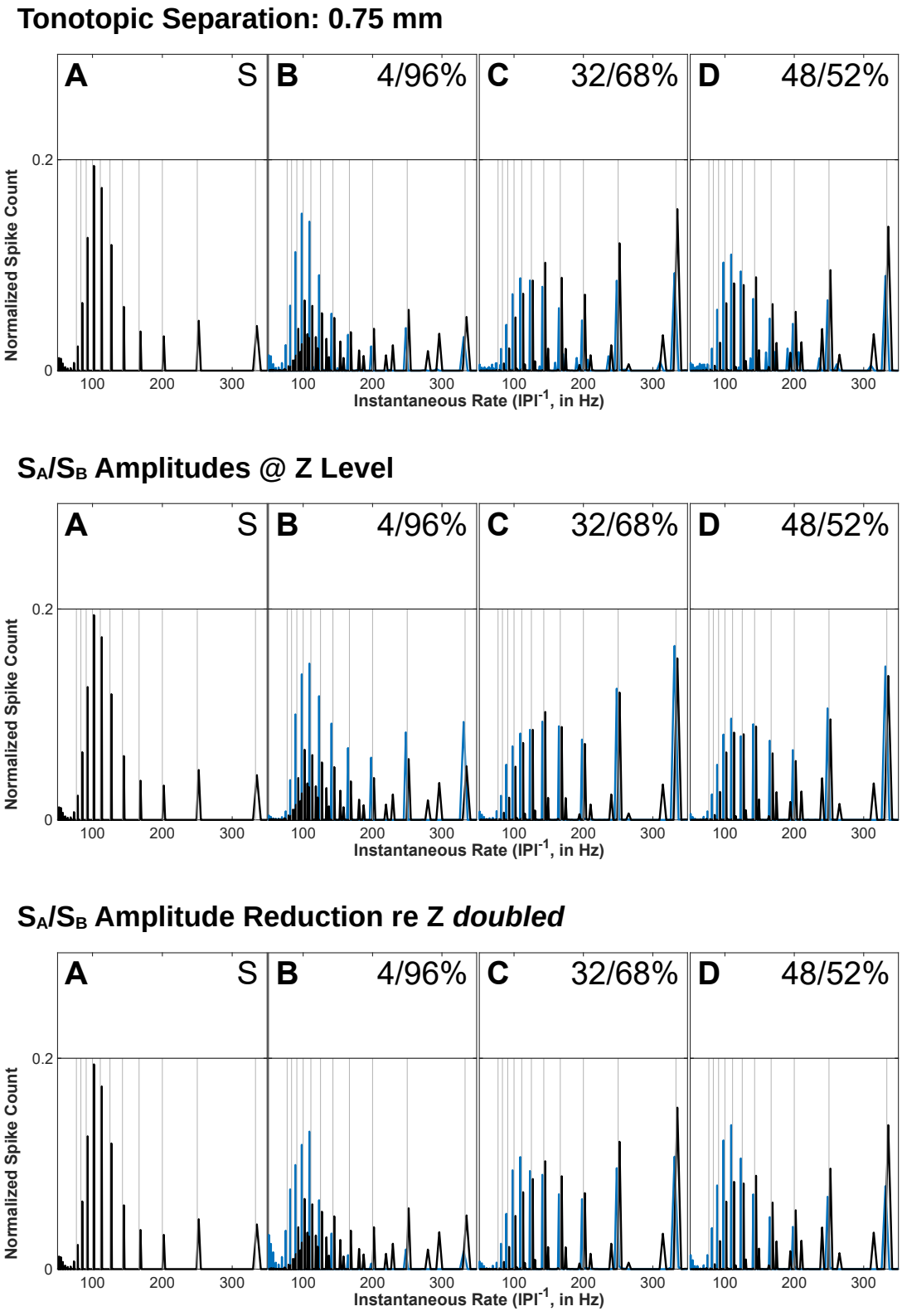


Figure S5

*Spike counts as a function of instantaneous rate (i.e., IPI^{-1}, in Hz) for **HR** with three variations of model parameters as compared to the psychophysical experiment setup (cf. Table 4). All other as-pects as in Figure 8.*

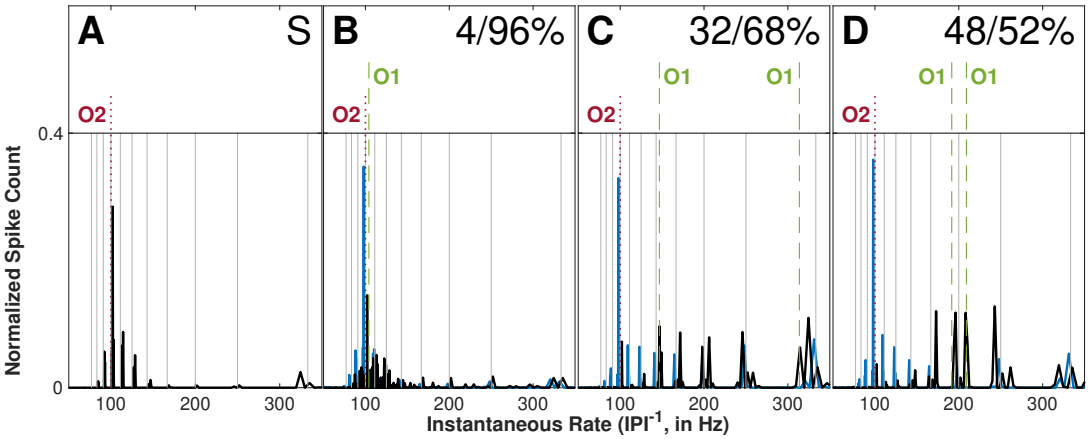


Figure S6

*Normalized spike counts as a function of instantaneous rate for **SIPI** stimuli. All other aspects as in Figure 7.*

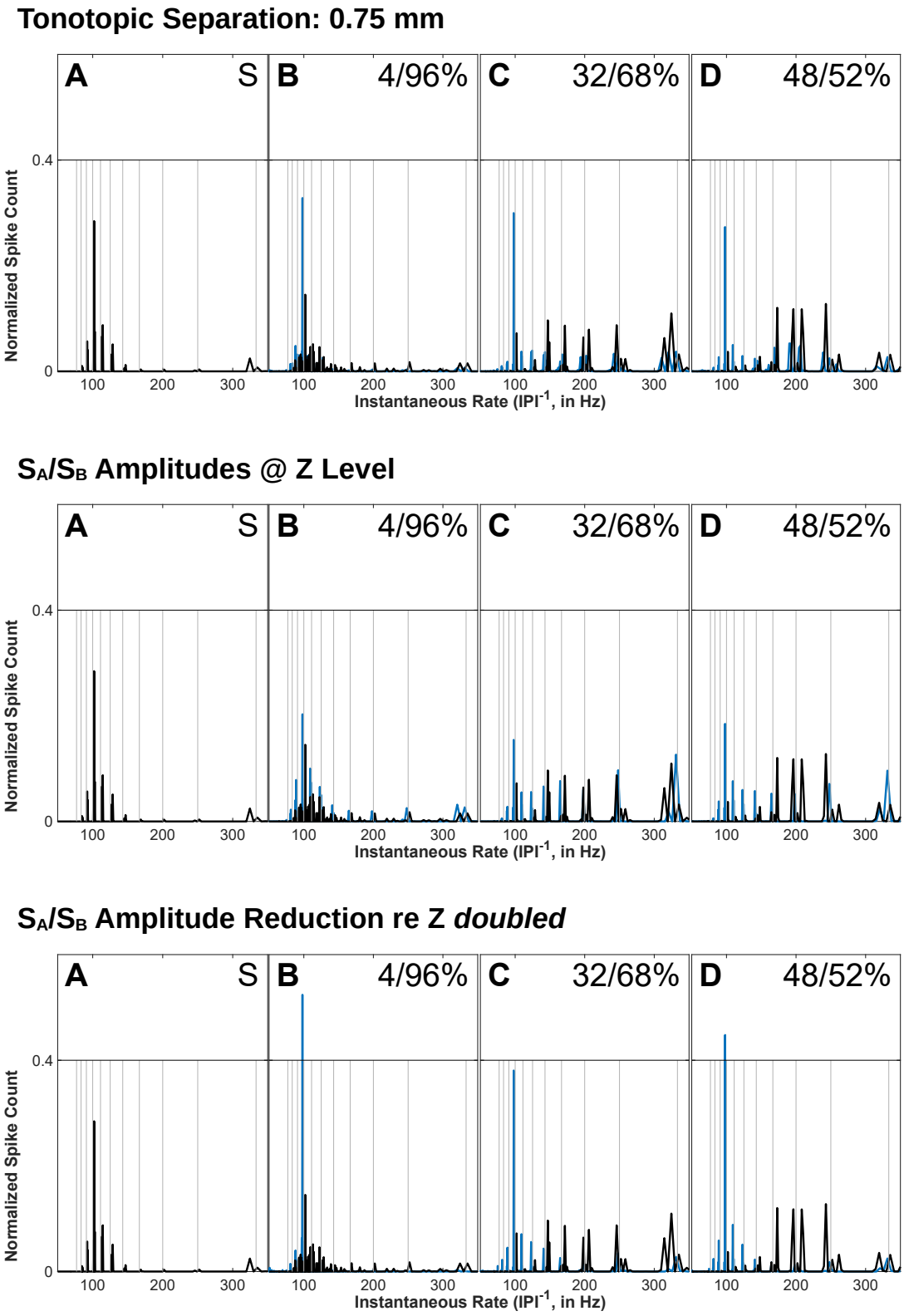


Figure S7

*Spike counts as a function of instantaneous rate (i.e., IPI^{-1}, in Hz) for **SIPI** with three variations of model parameters as compared to the psychophysical experiment setup (cf. Table 4). All other as-pects as in Figure 9.*

General Discussion

In this doctoral thesis, we investigated basic approaches to enhance timing-cue perception in CI listeners towards improving selective hearing. In the following, the main findings are discussed in a broader context, and future research directions are motivated to ultimately enhance electric hearing not only in the laboratory but also in the real world.

Various Temporal-Pitch Cues with Single-Electrode Stimulation

The first major aspect of timing-cue perception investigated in this thesis was the stimulation approach, that is, the principal means of encoding timing either on a pulse-by-pulse basis (LR stimulation), via the AM of constant-rate pulse trains (HR stimuli), or with modified AM stimuli involving SIPI pulses inserted at AM peaks (SIPI stimuli). All approaches were tested at a comfortable level.

Across studies, we found that single-electrode LR stimulation yields the highest timing-cue sensitivity. HR stimulation approaches (or, rather, is not significantly different from) this level of sensitivity only for large AM depths. SIPI stimulation approaches LR sensitivity for all AM depths, thereby increasing sensitivity for small AM depths most relevant in everyday scenarios. Although these results may suggest that—at least under simplified single-electrode conditions—SIPIs are a straightforward means to reduce the trade-off between HR stimulation (favoring speech cues) and LR stimulation (favoring timing cues), the comparative rather than absolute nature of our results merits further discussion in terms of whether SIPI and particularly LR sensitivity are high enough to provide relevant information in everyday listening situations.

To better assess the potential of temporal pitch with CIs, researchers have identified acoustic stimuli yielding similar sensitivity in NH listeners and in CI listeners when employing LR stimulation. The NH stimuli were found to be acoustic pulse trains bandpass-filtered in a medium to high frequency region (Carlyon & Deeks, 2002; Carlyon et al., 2002, 2008; van Wieringen et al., 2003). For such acoustic stimuli, the harmonics in the bandpass range are assumed to be unresolved (J. G. Bernstein & Oxenham, 2003). Hence, place pitch cues are assumed to be absent for such stimuli and listeners in turn rely purely on temporal-envelope

cues. Such temporal-envelope pitch is assumed to be weaker than pitch derived from resolved harmonics (which may provide temporal fine-structure, temporal-envelope, and place cues). For a review on NH pitch perception see, for example, Oxenham (2012).

To compare temporal-envelope pitch between CI and NH listeners, the upper limit of temporal pitch has received much attention. While NH studies suggest a limit of about 600 to 900 pps (Carlyon & Deeks, 2002; Macherey & Carlyon, 2014), CI studies vary in their conclusion but generally suggest an upper limit of 300 to 500 pps (De Groote et al., 2024; Ihlefeld et al., 2015; Kong et al., 2009; Lindenbeck, Majdak, et al., 2023; Lindenbeck et al., 2020; Townshend et al., 1987). However, star CI listeners show upper limits similar to NH listeners (e.g., Kong and Carlyon, 2010), suggesting that CI pitch is degraded beyond the principal limitations of purely temporal pitch by CI-specific factors. With the major difference between purely temporal NH and CI pitch being the health of the AN, researchers hypothesized that factors related to etiology, such as the duration of deafness, may explain the gap between average and maximum possible upper limit. However, while dedicated studies are lacking, a recent meta-analysis incorporating six studies with in total 50 CI listeners did not show a within-study effect of duration of deafness on the upper limit (Carlyon et al., 2025).

Thus, taken together, besides the conclusion that both SIPI and LR pitch cues might never be as salient as pitch cues from resolved harmonics conventionally available in NH, pitch perception of many CI listeners seems to fall short of the theoretically possible maximum with electric stimulation. The reasons for that remain largely unknown and are a crucial field for future research, as highlighted by the recent opinion paper on the past, present, and future of research on temporal processing with CIs (Carlyon et al., 2025).

Across all four studies incorporated in this thesis, we used two-interval two-alternative forced-choice tasks to measure basic discrimination sensitivity. However, in everyday selective-hearing tasks, acoustic input not only has to be discriminated but rather mentally segregated to form representations of the constituting physical sources (i.e., auditory objects). Furthermore, subsequently, these sources have to be tracked over time and through auditory space (i.e., they have to be streamed). For reviews on auditory grouping and stream segregation see, for example, Darwin (1997) and Carlyon (2004).

The timing cues focused on in this thesis provide potent information for both auditory object formation (temporal pitch) and streaming (ITD). Hence, to better grasp potential benefits of having access to salient timing cues in tasks requiring object formation and streaming, we conducted a study (not as part of this thesis) to test the rhythmic masking release (RMR) paradigm (Middlebrooks & Onsan, 2012) in CI listeners and with single-electrode LR stimulation only. RMR tests voluntary stream segregation with sequential presentation of competing streams in a rhythm discrimination task. For FDs between 6 and 27% and ITDs between 500 and 2000 µs, preliminary analyses (Lindenbeck, Frohmann, et al., 2023) suggest that RMR is possible based on pitch cues but not ITD cues when presented in isolation. When presenting both cues together, however, RMR performance was super-additive (i.e., synergistic). Taken together, the tentative results suggest that, with CIs, ITD alone may not be sufficient for sound segregation (cf. Darwin, 1997) but having access to both temporal-pitch and ITD cues may be crucial for improving selective hearing in the future.

Channel Interactions with Dual-Electrode Stimulation

The second major aspect investigated in this thesis was the effect of the across-electrode delay on timing-cue sensitivity when switching from single-electrode to dual-electrode stimulation (Lindenbeck et al., 2024, 2025). Dual-electrode stimulation was tested as a controlled step towards multi-electrode stimulation necessary for speech understanding. We found that increasing the delay (called monaural temporal electrode asynchrony or mTEA) from quasi-simultaneous stimulation towards half of the dominant stimulus period (i.e., the pulse period for LR stimuli and the AM/SIPI period for HR/SIPI stimuli) had a detrimental effect on timing-cue sensitivity. Furthermore, we found that this effect was similar for all three tested stimulus types and, for LR stimulation only, similar for unilateral (temporal pitch) and bilateral (ITD) stimulation.

To better understand the cues provided particularly by the more complex HR and SIPI stimuli, we simulated AN fiber responses to single-electrode stimuli and dual-electrode stimuli with electrode separations of zero and one electrode, respectively (cf. Lindenbeck et al., 2025). We repeatedly simulated a single AN fiber located on-channel at electrode six

(i.e., the C1 electrode in Lindenbeck et al. (2024, 2025)). Qualitative comparison between AN spiking and behavioral sensitivity patterns revealed that dual-electrode LR cues were altered for larger delays compared to single-electrode stimulation due to increases in effective rate as a result of channel interactions. Furthermore, dual-electrode HR cues were negatively affected by the simulated (and psychophysically required) reduction in per-electrode amplitude, likely reflecting broader dual-electrode AN excitation and, in turn, counteracted loudness integration (all simulations used loudness-balanced amplitudes from experiments). Similar to the behavioral results, modeled dual-electrode SIPI cues fell between LR and HR cues.

While there are numerous ways to further refine the modeling approach in future studies (e.g., aiming for quantitative prediction of the behavioral data, thereby possibly also modeling sub-cortical structures and cortical decision stages), we took the closest possible next step and modeled a simplified 2D model of the electrically stimulated cochlea (cf. Lindenbeck and Laback, 2025), largely following the approach of Takanen et al. (2024). Peri-stimulus time histograms (PSTHs) of AN fibers in response to LR, HR, and SIPI pulse trains presented at electrodes five and six of a fictional MED-EL FLEX28 array with a dominant rate of 100 Hz are shown in Figure I for an mTEA of 5% or 500 µs, respectively, and in Figure II for an mTEA of 35% or 3500 µs, respectively.[4]

In both figures, the gray horizontal ribbon marks the AN fibers lying between the two stimulating electrodes. These fibers are most strongly stimulated by both electrodes and, hence, are most prone to channel interactions manifesting as a mixture of the two per-electrode stimuli within a single fiber. In contrast, the fibers lying in the white areas are dominantly stimulated by the more proximal of the two electrodes and, hence, should largely reflect the stimulus emitted from the proximal electrode in their firing patterns. To discuss the differences in neural periodicity coding between single-electrode and dual-electrode stimulation in this still simplified scenario, comparing firing patterns in the gray and white areas might provide hints towards favorable stimulus configurations with respect to mTEA and timing-cue sensitivity (which crucially relies on genuine encoding of stimulus periodicity).

[4] Note that, compared to Lindenbeck et al. (2024, 2025), the mTEAs were slightly changed so that the compound carrier pulse pattern was regular (i.e., the mTEA [in µs] was $k + 0.5$ times the carrier pulse period [$k \in \mathbb{N}_0$]).

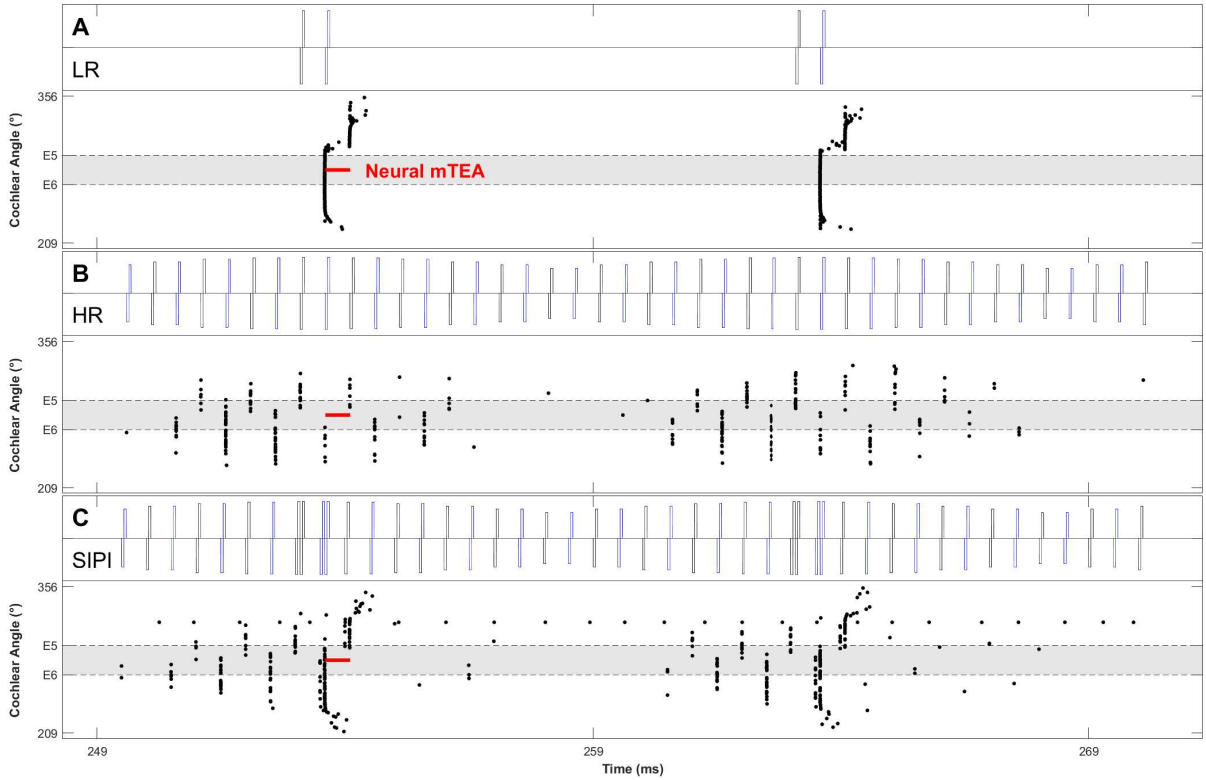


Figure I

Peri-stimulus time histograms of simulated AN fibers (Takanen & Seeber, 2022) in response to dual-electrode stimulation (i.e., condition S_A in Lindenbeck et al. (2024, 2025)) with a dominant rate of 100 Hz. The simulated setup was similar to that in Lindenbeck et al. (2025) except for that a simplified 2D model (cf. Takanen et al., 2024) was additionally implemented. The ordinate depicts the insertion angle as a cochlear-size-normalized proxy of electrode location. Stimulus types are shown in separate panels. The ***across-electrode delay*** *(i.e., the mTEA) was always 5% of the dominant periodicity, that is,* ***500 µs****. The more basal electrode 6 is always leading. The "neural mTEA" is marked with a red horizontal line. In each panel, the stimulus is depicted at the top (electrode 5 in blue and electrode 6 in black). These results were reported in Lindenbeck and Laback (2025).*

For small mTEAs in LR stimulation (cf. Figure I A), firing patterns are similar in the gray and white areas of the PSTHs, indicating that single-electrode spiking is largely preserved for dual-electrode stimulation. In other words, the physical mTEA is not reflected neurally (i.e., no spiking occurs after the neural mTEA as indicated in the PSTH with red horizontal lines). This parallels the behavioral sensitivity for this condition (cf. Figure 5A in Lindenbeck et al. (2025)). The simulated AN fibers in the gray ribbon that almost uniformly respond to the pulse emitted from electrode six are then still in a refractory state when the delayed pulse is emitted from the more apical electrode five.

Compared to that, for HR stimulation (cf. Figure I B), more spikes are triggered in the gray ribbon compared to the white outer areas. Although often a single neuron might not fire in response to subsequent pulses emitted from different electrodes, assuming that spikes from many neurons may be integrated into a compound code upstream, the effective rate for HR stimuli likely increases when switching from single-electrode to dual-electrode stimulation. However, since HR-based pitch is not coded pulse by pulse but via the AM period, the small mTEA likely results in little changes in behavioral sensitivity as it affects the effective AM depth only little and preserves AM periodicity in the compound spike pattern. This line of reasoning parallels the behavioral HR results well (cf. Figure 5B of Lindenbeck et al. (2025)).

Similar to the behavioral SIPI results reported throughout this thesis, also the SIPI PSTH (cf. Figure I C) combines properties of LR and HR PSTHs. The SIPI pulses trigger spiking patterns similar to LR stimulation while the non-SIPI carrier pulses at the AM onset trigger spiking patterns similar to HR stimulation. Importantly, the suppression of the AM offset also occurs reliably in the AN fiber in the grey ribbon, thus, likely also improving periodicity coding by lowering the effective within-fiber rate with dual-electrode (or possibly multi-electrode) stimulation.

Increasing the mTEA from 500 to 3500 µs has a crucial effect on the periodicity encoded in the AN spiking patterns (cf. Figure II). Focusing on AN fibers the gray ribbons, for LR stimulation (panel A), pulses from both electrodes now elicit spikes within the same AM fiber, thus, manifesting channel interactions by increasing the effective rate and, as a consequence, likely reducing the salience of the encoded pulse period. Following Carlyon et al. (2002), the perceived pitch is largely determined by the larger of the two inter-pulse (or rather inter-spike) intervals, that is, 6500 µs or 154 pps. As for the small mTEA, the mTEA-induced changes in spiking patterns are similar for SIPI stimulation (panel C) and LR stimulation. Notably, spiking in response to either LR or SIPI pulses from both electrodes within the same AN fibers introduces a neural mTEA (marked with red horizontal lines) related neither to the AM period (i.e., the desired periodicity) nor to the carrier pulse period. For HR stimulation (panel B), the mTEA results in a reduction in neural AM depth which manifests as a reduced neural off-time (i.e., continuous spiking throughout the per-electrode

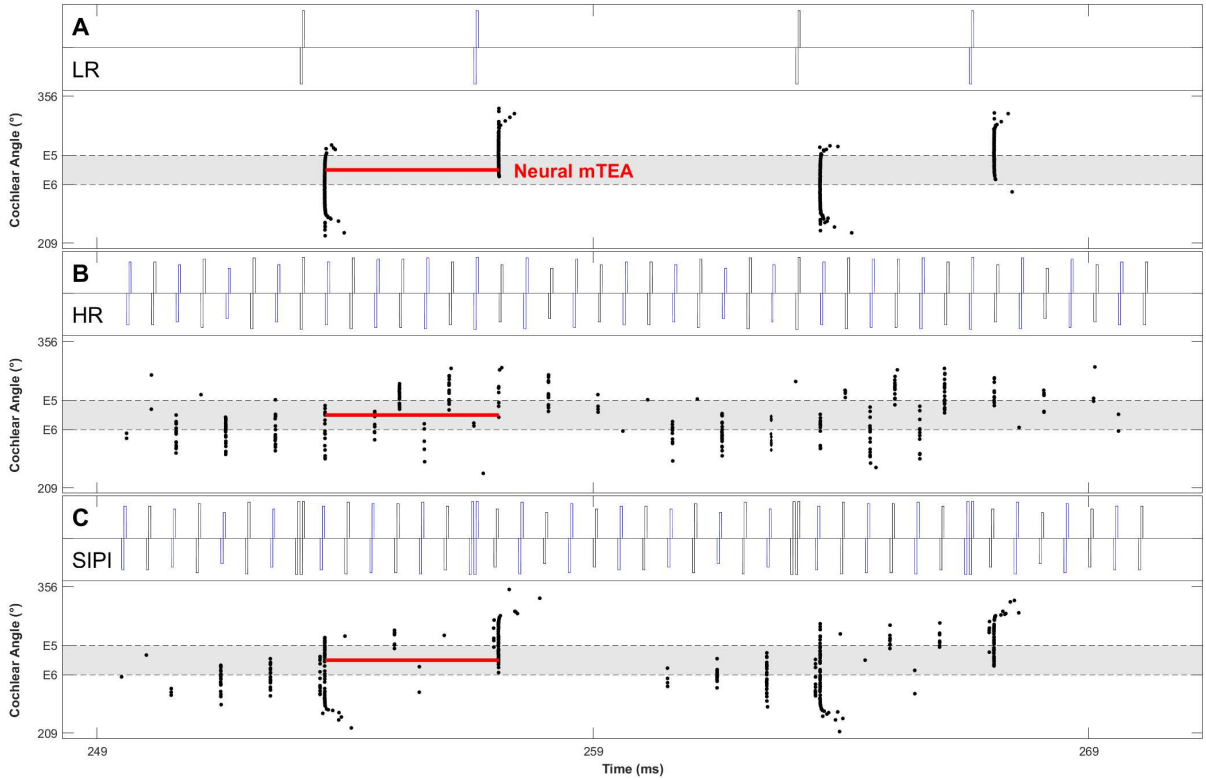


Figure II
*Peri-stimulus time histograms of simulated AN fibers in response to dual-electrode stimulation with a **between-electrode delay** (mTEA) of 35% or **3500 µs**. All other aspects as in Figure I.*

AM valleys) compared to the small mTEA.

Taken together, the behavioral results gathered in this thesis with dual-electrode stimulation in combination with the phenomenological modeling of AN spiking patterns in response to these stimuli suggests that employing small mTEAs in multi-electrode stimulation can at least somewhat preserve the per-electrode coding either due to the exploitation of refractory effects or due to preservation of the AM depth. Furthermore, assuming that high-rate stimulation is necessary to reliably encode speech (Arora et al., 2009; Friesen et al., 2005; Loizou et al., 2000), SIPI stimulation may be an effective means to introduce the preciseness of LR-induced spiking into HR stimulation patterns. Furthermore, extended future modeling studies may evaluate the potential of SIPI stimulation to “design“ channel interactions by carefully timing SIPIs across electrodes and, thus, finding the optimal combination of SIPI rate and mTEA in connection with a given AM rate.

Towards Improving Selective Hearing: Future Directions and Outlook

This thesis provides several basic findings, in particular regarding SIPI stimulation, which—beyond a better understanding of timing-cue perception with CIs—may be used to improve electric hearing with future devices incorporating the findings into their stimulation strategies. However, several crucial further steps are required for any SIPI-based benefit to manifest maximally in everyday CI listening.

First, a multi-electrode stimulation strategy incorporating SIPIs has to be designed and evaluated. At the beginning of this doctoral project, we sketched a crude paradigm called *MissiSIPI* (Lindenbeck et al., 2019) based on an FS4 processing scheme (e.g., Riss et al., 2014) which, however, was not further refined during the course of the project.

Recently, Fan et al. (2025) presented pilot data from three bilateral Advanced Bionics CI listeners. They compared bilateral 16-electrode SIPI stimulation, motivated by our earlier studies, with CIS stimulation in two tasks. ITD-based left/right discrimination of 500 Hz pure tones and narrowband noise (ITDs < 600 µs) was at chance level with CIS, yet, with SIPI rates between 50 and 200 pps, discrimination approached threshold levels (i.e., scores of approximately 70%) for SIPI rates of 50 and 100 pps, respectively. Notably, Fan et al. (2025) closely followed the approach of applying SIPIs roughly at integer sub-multiples of the dominant periodicity (here, the AM rate), as suggested by Lindenbeck, Majdak, et al. (2023) and by Srinivasan et al. (2020). Furthermore, at 0 dB signal-to-noise ratio, phoneme identification in noise was similar between CIS and 50-pps SIPI stimulation in the co-located N_0S_0 condition but about 10% better for SIPI stimulation in the N_0S_π condition. Taken together, these pilot results suggest that multi-channel SIPI stimulation may indeed improve spatial hearing and speech-in-noise perception without sacrificing basic speech understanding.

However, since Fan et al. (2025) did not report on the employed mTEAs and furthermore did not compare their approach to mixed-rate (e.g., Riss et al., 2014; Thakkar et al., 2018) or even all-low-rate strategies (e.g., Smith, 2010), it is unclear whether their approach provided the maximum possible timing-cue sensitivity. Instead, it might be further improved by (1) optimizing mTEAs or (2) introducing electrodes with LR-like stimulation. In

any case, the *asymmetric* SIPI-induced AM enhancement which simultaneously enhances AM peaks and suppresses AM offsets (cf. Appendix of Lindenbeck et al. (2025)) may be incorporated into a SIPI-based stimulation strategy by canceling several post-SIPI carrier pulses and thereby reducing both channel interaction potential and CI energy consumption.

Second, any promising SIPI-based stimulation strategy requires electrode arrays that target AN fibers as optimally as possible (i.e., it requires optimized ENIs). Different CI manufacturers follow different design approaches derived from different weighting of the major design dimensions (1) proximity of the electrodes to the AN, (2) preservation of mechanical cochlear structures, and (3) cochlear coverage and stimulation of apical AN fibers (e.g., Dhanasingh et al., 2024; Risi, 2019). On the one hand, favoring (1) over (2) and (3), pre-curved (stiff) perimodiolar (PM) arrays were designed. On the other hand, favoring (2) and (3) over (1), straight (flexible) lateral-wall (LW) arrays were designed. Within manufacturer, PM arrays provide higher speech understanding than LW arrays (e.g., Sharma et al., 2022) and long LW arrays provide higher speech understanding than short LW arrays (e.g., Canfarotta et al., 2021). Across manufacturers and when controlling for confounding factors (e.g., duration of deafness, pre-operative audiometry), differences between manufacturers in speech understanding are small, if present at all (e.g., Fabie et al., 2018; Haumann et al., 2010).

Beyond speech outcomes, further support for the emphasis on full cochlear coverage and apical stimulation comes from animal studies. In anesthetized cats, Middlebrooks and Snyder (2010) found a pathway with particularly high temporal acuity in apical AN fibers. This is in line with recent results on basilar-membrane mechanics (Burwood et al., 2022) suggesting that, in guinea pigs, place specificity is absent in the cochlear apex. Higher temporal acuity in apical AN fibers could motivate CI manufacturers to increase the array insertion depth to reach as many apical fibers as possible. Currently, only MED-EL follows the design principle of “ deep“ two-turn (i.e., 720 degree) insertion. Studies found better apical rate discrimination (e.g., Stahl et al., 2016) but similar upper limits of temporal pitch (e.g., De Groote et al., 2024). This could be attributed to the insufficient ENI obtained with monopolar stimulation used in MED-EL devices in terms of current flow (Carlyon et al., 2025; De Groote et al., 2024) or a generally insufficient insertion depth.

Support for the argument of suboptimal current flow with monopolar stimulation comes from a recent pilot study placing a ground electrode in the cochlear helicotrema via an apical cochleostomy (Landsberger et al., 2022) to complement a short electrode array manufactured by Cochlear Corp. Pilot results showed a lower perceived pitch with the apical ground as compared to regular monopolar stimulation. However, while the current flow might be optimized for apical electrodes, potential negative effects of changes in current flow for basal electrodes (which are further distant from the apical ground as compared to monopolar stimulation) particularly with respect to speech understanding remain to be excluded.

Support for the argument of insufficient insertion depth comes from recent studies showing that even with 31.5 mm MED-EL arrays, only 92% of the intended two cochlear turns are covered (Dhanasingh et al., 2024). The coverage is with 80% even worse for the increasingly popular 28 mm MED-EL FLEX28 array—which is designed to provide the same two-turn insertion for smaller cochleae—suggesting that surgeons may sometimes opt for the shorter but easier to implant array rather than the longest possible array, therewith sacrificing the possibility of "truly" apical stimulation. Results such as those by Breitsprecher et al. (2022), showing that most methods to assess the cochlear duct length (which is required to select the longest possible electrode array) underestimate the actual length, further extend the list of reasons for insufficient insertion depths. Reacting to the potential underestimation of cochlear duct length, MED-EL has recently introduced an even longer 34 mm FLEX34 array (for details see, e.g., MED-EL, 2024). Initial reports show that every sixth to tenth adult in two major Austrian implantation clinics is eligible for the FLEX34 array (Baumgartner, 2024). Hence, while comprehensive analyses with this array are still lacking, CI listeners with such long arrays may be a promising cohort for future studies on human apical timing-cue sensitivity and, in turn, potential selective-hearing benefits.

Third, any promising SIPI-based stimulation strategy operating on an optimized ENI will not unleash its maximum potential without training and extended daily usage. With respect to controlled training, Goldsworthy and Shannon (2014) have shown that single-electrode rate discrimination improved by more than a factor of two after 32 hours of training. With even more extensive training, Raymond L. Goldsworthy—who is a CI

listener—has recently reported in Carlyon et al. (2025) to have a self-estimated upper limit of temporal pitch of about 880 pps (cf. Kong and Carlyon, 2010), with a weak percept extending up to the technical device limit of 3520 pps. These reports parallel our anecdotal experiences in the laboratory with some CI listeners with extensive music training (e.g., CI18 who participated in Lindenbeck, Majdak, et al. (2023) and Lindenbeck et al. (2020)) having lower-than-reported rate discrimination thresholds of about 1% or even less (cf. Carlyon et al., 2025; Moore and Carlyon, 2005). Training studies focusing on one narrow aspect of perception (e.g., rate pitch) have been questioned as they may not genuinely reflect sensory plasticity (e.g., Carlyon et al., 2025) and hence, may generalize only poorly to other related aspects of perception. Towards training more complex perception, we started a project on musical harmony training with initial findings revealing fortunate stimulus configurations for high chord discrimination sensitivity with clinical CI processors (Augsten et al., 2023, 2025a) as well as behavioral and neural sensitivity to harmonically relevant differences in pitch *relations* (Augsten et al., 2025b)—so far, however, without controlled training.

In addition to focused training conducted under controlled (or even laboratory) conditions, regular real-world CI usage is crucial for fast habituation and comprehensive training. Among many other technical factors, such regular usage may be prevented by social stigma attributed to wearing visible prostheses (e.g., Wallhagen, 2010). Hence, making CI processors smaller or even invisible may be a booster for CI acceptance and a larger proportion of CI candidates being actually supplied. Towards this goal, the totally implantable CI (TICI) was developed and recently successfully tested in six human CI listeners (Lefebvre et al., 2025). Initial analyses showed speech understanding both in quiet and noise to be at the level of conventional CIs. Furthermore, the TICI was well accepted and used throughout the day.

Taken together, all these results inspire optimism that a holistic approach combining insights from basic to applied research and medicine can ultimately reduce unsupplied hearing loss. Currently, more than 90% of CI candidates worldwide are still unsupplied (cf. Lefebvre et al., 2025). Hence, there is still a long way to go towards the ultimate goal of providing every eligible and consenting person with a hearing prosthesis that largely restores complex and, crucially, selective hearing.

This page is intentionally left blank

References

Arora, K., Dawson, P., Dowell, R., & Vandali, A. (2009). Electrical stimulation rate effects on speech perception in cochlear implants. *International Journal of Audiology*, *48*, 561–567. https://doi.org/10.1080/14992020902858967

Augsten, M.-L., Lindenbeck, M. J., & Laback, B. (2023). *Perception of Simultaneous and Sequential Musical Harmony in Cochlear Implant Listeners* [Presentation]. Conference on Implantable Auditory Prostheses, Lake Tahoe, CA, USA.

Augsten, M.-L., Lindenbeck, M. J., & Laback, B. (2025a). Discrimination of spectrally sparse musical chords in cochlear implant listeners [manuscript submitted for publication]. *Acoustics Research Institute, Austrian Academy of Sciences*.

Augsten, M.-L., Lindenbeck, M. J., & Laback, B. (2025b). *Neural Processing of Pitch Relations within Spectrally Sparse Musical Chords in Cochlear Implant Listeners* [Presentation]. DAS | DAGA 2025 - 51st Annual Meeting on Acoustics, Copenhagen, Denmark.

Baumgartner, W.-D. (2024). Die neue Flex-34 Cochlea Implantat Elektrode [The new Flex-34 Cochlear Implant Electrode]. *Wiener Medizinische Wochenzeitschrift Skriptum*, *21*(4), 15–16.

Bernstein, J. G., & Oxenham, A. J. (2003). Pitch discrimination of diotic and dichotic tone complexes: Harmonic resolvability or harmonic number? *The Journal of the Acoustical Society of America*, *113*(6), 3323. https://doi.org/10.1121/1.1572146

Bernstein, J. G., Jensen, K. K., Stakhovskaya, O. A., Noble, J. H., Hoa, M., Kim, H. J., Shih, R., Kolberg, E., Cleary, M., & Goupell, M. J. (2021). Interaural Place-of-Stimulation Mismatch Estimates Using CT Scans and Binaural Perception, But Not Pitch, Are Consistent in Cochlear-Implant Users. *The Journal of Neuroscience*, *41*(49), 10161–10178. https://doi.org/10.1523/JNEUROSCI.0359-21.2021

Bernstein, L. R., & Trahiotis, C. (2002). Enhancing sensitivity to interaural delays at high frequencies by using “transposed stimuli”. *The Journal of the Acoustical Society of America*, *112*(3), 1026–1036. https://doi.org/10.1121/1.1497620

Boulet, J., White, M., & Bruce, I. C. (2016). Temporal considerations for stimulating spiral ganglion neurons with cochlear implants. *Journal of the Association for Research in Otolaryngology*, *17*(1), 1–17. https://doi.org/10.1007/s10162-015-0545-5

Breitsprecher, T., Dhanasingh, A., Schulze, M., Kipp, M., Dakah, R. A., Oberhoffner, T., Dau, M., Frerich, B., Weber, M.-A., Langner, S., Mlynski, R., & Weiss, N. M. (2022). CT imaging-based approaches to cochlear duct length estimation—a human temporal bone study. *European Radiology*, *32*(2), 1014–1023. https://doi.org/10.1007/s00330-021-08189-x

Bronkhorst, A. W., & Plomp, R. (1988). The effect of head-induced interaural time and level differences on speech intelligibility in noise. *The Journal of the Acoustical Society of America*, *83*(4), 1508–1516. https://doi.org/10.1121/1.395906

Brughera, A., Dunai, L., & Hartmann, W. M. (2013). Human interaural time difference thresholds for sine tones: The high-frequency limit. *The Journal of the Acoustical Society of America*, *133*(5), 2839–2855. https://doi.org/10.1121/1.4795778

Buechel, B. D., Hancock, K. E., Chung, Y., & Delgutte, B. (2018). Improved neural coding of ITD with bilateral cochlear implants by introducing short inter-pulse intervals. *Journal of the Association for Research in Otolaryngology*, *19*(6), 681–702. https://doi.org/10.1007/s10162-018-00693-0

Burwood, G., Hakizimana, P., Nuttall, A. L., & Fridberger, A. (2022). Best frequencies and temporal delays are similar across the low-frequency regions of the guinea pig cochlea. *Science Advances*, *8*(38), eabq2773. https://doi.org/10.1126/sciadv.abq2773

Canfarotta, M. W., Dillon, M. T., Buchman, C. A., Buss, E., O'Connell, B. P., Rooth, M. A., King, E. R., Pillsbury, H. C., Adunka, O. F., & Brown, K. D. (2021). Long-Term Influence of Electrode Array Length on Speech Recognition in Cochlear Implant Users. *The Laryngoscope*, *131*(4), 892–897. https://doi.org/10.1002/lary.28949

Carlyon, R. P. (2004). How the brain separates sounds. *Trends in Cognitive Sciences*, *8*(10), 465–471. https://doi.org/10.1016/j.tics.2004.08.008

Carlyon, R. P., & Deeks, J. M. (2002). Limitations on rate discrimination. *The Journal of the Acoustical Society of America*, *112*(3), 1009–1025. https://doi.org/10.1121/1.1496766

Carlyon, R. P., Deeks, J. M., Delgutte, B., Chung, Y., Vollmer, M., Ohl, F. W., Kral, A., Tillein, J., Litovsky, R. Y., Schnupp, J., Rosskothen-Kuhl, N., & Goldsworthy, R. L. (2025). Limitations on Temporal Processing by Cochlear Implant Users: A Compilation of Viewpoints. *Trends in Hearing*, *29*, 23312165251317006. https://doi.org/10.1177/23312165251317006

Carlyon, R. P., Long, C. J., & Deeks, J. M. (2008). Pulse-rate discrimination by cochlear-implant and normal-hearing listeners with and without binaural cues. *The Journal of the Acoustical Society of America*, *123*(4), 2276–2286. https://doi.org/10.1121/1.2874796

Carlyon, R. P., van Wieringen, A., Long, C. J., Deeks, J. M., & Wouters, J. (2002). Temporal pitch mechanisms in acoustic and electric hearing. *The Journal of the Acoustical Society of America*, *112*(2), 621–633. https://doi.org/10.1121/1.1488660

Cartee, L. A., Miller, C. A., & van den Honert, C. (2006). Spiral ganglion cell site of excitation I: Comparison of scala tympani and intrameatal electrode responses. *Hearing Research*, *215*(1–2), 10–21. https://doi.org/10.1016/j.heares.2006.02.012

Cartee, L. A., van den Honert, C., Finley, C. C., & Miller, R. L. (2000). Evaluation of a model of the cochlear neural membrane. I. Physiological measurement of membrane characteristics in response to intrameatal electrical stimulation. *Hearing Research*, *146*(1–2), 143–152. https://doi.org/10.1016/S0378-5955(00)00109-X

Chatterjee, M., & Oberzut, C. (2011). Detection and rate discrimination of amplitude modulation in electrical hearing. *The Journal of the Acoustical Society of America*, *130*(3), 1567–1580. https://doi.org/10.1121/1.3621445

Croghan, N. B. H., Duran, S. I., & Smith, Z. M. (2017). Re-examining the relationship between number of cochlear implant channels and maximal speech intelligibility. *The Journal of the Acoustical Society of America*, *142*(6), EL537–EL543. https://doi.org/10.1121/1.5016044

Darwin, C. J. (1997). Auditory grouping. *Trends in Cognitive Sciences*, *1*(9), 327–333. https://doi.org/10.1016/S1364-6613(97)01097-8

de Balthasar, C., Boëx, C., Cosendai, G., Valentini, G., Sigrist, A., & Pelizzone, M. (2003). Channel interactions with high-rate biphasic electrical stimulation in cochlear implant subjects. *Hearing Research*, *182*(1–2), 77–87. https://doi.org/10.1016/S0378-5955(03)00174-6

De Groote, E., Carlyon, R. P., Deeks, J. M., & Macherey, O. (2024). Effects of selective stimulation of apical electrodes on temporal pitch perception by cochlear implant recipients. *The Journal of the Acoustical Society of America*, *156*(3), 2060–2076. https://doi.org/10.1121/10.0029023

Dhanasingh, A., Nielsen, S. B., Beal, F., Schilp, S., Hessler, R., Jolly, C., & Hochmair, I. (2024). Cochlear implant electrode design for safe and effective treatment. *Frontiers in Neurology*, *15*, 1348439. https://doi.org/10.3389/fneur.2024.1348439

Dietz, M., Marquardt, T., Salminen, N. H., & McAlpine, D. (2013). Emphasis of spatial cues in the temporal fine structure during the rising segments of amplitude-modulated sounds. *Proceedings of the National Academy of Sciences*, *110*(37), 15151–15156. https://doi.org/10.1073/pnas.1309712110

Dillon, M. T., Buss, E., King, E. R., Deres, E. J., Obarowski, S. N., Anderson, M. L., & Adunka, M. C. (2016). Comparison of two cochlear implant coding strategies on speech perception. *Cochlear Implants International*, *17*(6), 263–270. https://doi.org/10.1080/14670100.2016.1244033

Dynes, S. B. C., & Delgutte, B. (1992). Phase-locking of auditory-nerve discharges to sinusoidal electric stimulation of the cochlea. *Hearing Research*, *58*(1), 79–90. https://doi.org/10.1016/0378-5955(92)90011-b

Egger, K., Majdak, P., & Laback, B. (2016). Channel interaction and current level affect across-electrode integration of interaural time differences in bilateral cochlear-implant listeners. *Journal of the Association for Research in Otolaryngology*, *17*(1), 55–67. https://doi.org/10.1007/s10162-015-0542-8

Fabie, J. E., Keller, R. G., Hatch, J. L., Holcomb, M. A., Camposeo, E. L., Lambert, P. R., Meyer, T. A., & McRackan, T. R. (2018). Evaluation of Outcome Variability Associated With Lateral Wall, Mid-scalar, and Perimodiolar Electrode Arrays When

Controlling for Preoperative Patient Characteristics. *Otology & Neurotology*, *39*(9), 1122–1128. https://doi.org/10.1097/MAO.0000000000001951

Fan, Y., Wu, Y., & Gifford, R. (2025). *Deploying Short Interpulse Interval (SIPI) Strategy in Multi-channel Systems* [Poster presentation]. 48th Annual MidWinter Meeting, Orlando, FL, USA.

Felsheim, R. C., & Dietz, M. (2024). An Adaptive Leaky-Integrate and Firing Probability Model of an Electrically Stimulated Auditory Nerve Fiber. *Trends in Hearing*, *28*, 23312165241286742. https://doi.org/10.1177/23312165241286742

Fischer, T., Schmid, C., Kompis, M., Mantokoudis, G., Caversaccio, M., & Wimmer, W. (2021). Effects of temporal fine structure preservation on spatial hearing in bilateral cochlear implant users. *The Journal of the Acoustical Society of America*, *150*(2), 673–686. https://doi.org/10.1121/10.0005732

Friesen, L. M., Shannon, R. V., Baskent, D., & Wang, X. (2001). Speech recognition in noise as a function of the number of spectral channels: Comparison of acoustic hearing and cochlear implants. *The Journal of the Acoustical Society of America*, *110*(2), 1150–1163. https://doi.org/10.1121/1.1381538

Friesen, L. M., Shannon, R. V., & Cruz, R. J. (2005). Effects of Stimulation Rate on Speech Recognition with Cochlear Implants. *Audiology and Neurotology*, *10*(3), 169–184. https://doi.org/10.1159/000084027

Fu, Q.-J., & Nogaki, G. (2005). Noise Susceptibility of Cochlear Implant Users: The Role of Spectral Resolution and Smearing. *Journal of the Association for Research in Otolaryngology*, *6*(1), 19–27. https://doi.org/10.1007/s10162-004-5024-3

Galvin, J. J., Oba, S., Başkent, D., & Fu, Q.-J. (2015). Modulation frequency discrimination with single and multiple channels in cochlear implant users. *Hearing Research*, *324*, 7–18. https://doi.org/10.1016/j.heares.2015.02.007

Gaudrain, E., & Carlyon, R. P. (2013). Using Zebra-speech to study sequential and simultaneous speech segregation in a cochlear-implant simulation. *The Journal of the Acoustical Society of America*, *133*(1), 502–518. https://doi.org/10.1121/1.4770243

Geurts, L., & Wouters, J. (2001). Coding of the fundamental frequency in continuous interleaved sampling processors for cochlear implants. *The Journal of the Acoustical Society of America*, *109*(2), 713–726. https://doi.org/10.1121/1.1340650

Goldsworthy, R. L., Bissmeyer, S. R. S., & Camarena, A. (2022). Advantages of Pulse Rate Compared to Modulation Frequency for Temporal Pitch Perception in Cochlear Implant Users. *Journal of the Association for Research in Otolaryngology*, *23*(1), 137–150. https://doi.org/10.1007/s10162-021-00828-w

Goldsworthy, R. L., & Shannon, R. V. (2014). Training improves cochlear implant rate discrimination on a psychophysical task. *The Journal of the Acoustical Society of America*, *135*(1), 334–341. https://doi.org/10.1121/1.4835735

Goupell, M. J., Kan, A., & Litovsky, R. Y. (2013). Mapping procedures can produce non-centered auditory images in bilateral cochlear implantees. *The Journal of the Acoustical Society of America*, *133*(2), EL101–EL107. https://doi.org/10.1121/1.4776772

Goupell, M. J., Stoelb, C., Kan, A., & Litovsky, R. Y. (2013). Effect of mismatched place-of-stimulation on the salience of binaural cues in conditions that simulate bilateral cochlear-implant listening. *The Journal of the Acoustical Society of America*, *133*(4), 2272–2287. https://doi.org/10.1121/1.4792936

Grantham, D. W., Ashmead, D. H., Ricketts, T. A., Haynes, D. S., & Labadie, R. F. (2008). Interaural Time and Level Difference Thresholds for Acoustically Presented Signals in Post-Lingually Deafened Adults Fitted with Bilateral Cochlear Implants Using CIS+ Processing. *Ear and Hearing*, *29*(1), 33–44. https://doi.org/10.1097/AUD.0b013e31815d636f

Graves, J. E., & Oxenham, A. J. (2019). Pitch discrimination with mixtures of three concurrent harmonic complexes. *The Journal of the Acoustical Society of America*, *145*(4), 2072–2083. https://doi.org/10.1121/1.5096639

Green, T., Faulkner, A., & Rosen, S. (2004). Enhancing temporal cues to voice pitch in continuous interleaved sampling cochlear implants. *The Journal of the Acoustical Society of America*, *116*(4), 2298–2310. https://doi.org/10.1121/1.1785611

Green, T., Faulkner, A., Rosen, S., & Macherey, O. (2005). Enhancement of temporal periodicity cues in cochlear implants: Effects on prosodic perception and vowel identification. *The Journal of the Acoustical Society of America*, *118*(1), 375–385. https://doi.org/10.1121/1.1925827

Hancock, K. E., Chung, Y., & Delgutte, B. (2012). Neural ITD coding with bilateral cochlear implants: Effect of binaurally coherent jitter. *Journal of Neurophysiology*, *108*(3), 714–728. https://doi.org/10.1152/jn.00269.2012

Hancock, K. E., Chung, Y., McKinney, M. F., & Delgutte, B. (2017). Temporal envelope coding by inferior colliculus neurons with cochlear implant stimulation. *Journal of the Association for Research in Otolaryngology*, *18*(6), 771–788. https://doi.org/10.1007/s10162-017-0638-4

Haumann, S., Lenarz, T., & Büchner, A. (2010). Speech Perception with Cochlear Implants as Measured Using a Roving-Level Adaptive Test Method. *ORL*, *72*(6), 312–318. https://doi.org/10.1159/000318872

Heffer, L. F., Sly, D. J., Fallon, J. B., White, M. W., Shepherd, R. K., & O'Leary, S. J. (2010). Examining the Auditory Nerve Fiber Response to High Rate Cochlear Implant Stimulation: Chronic Sensorineural Hearing Loss and Facilitation. *Journal of Neurophysiology*, *104*(6), 3124–3135. https://doi.org/10.1152/jn.00500.2010

Hochmair, I., Nopp, P., Jolly, C., Schmidt, M., ser, H. S. s., Garnham, C., & Anderson, I. (2006). MED-EL cochlear implants: State of the art and a glimpse into the future. *Trends in Amplification*, *10*(4), 201–219. https://doi.org/10.1177/1084713806296720

Horne, C. D. F., Sumner, C. J., & Seeber, B. U. (2016). A Phenomenological Model of the Electrically Stimulated Auditory Nerve Fiber: Temporal and Biphasic Response Properties. *Frontiers in Computational Neuroscience*, *10*. https://doi.org/10.3389/fncom.2016.00008

Hu, H., & Dietz, M. (2015). Comparison of Interaural Electrode Pairing Methods for Bilateral Cochlear Implants. *Trends in Hearing*, *19*, 1–22. https://doi.org/10.1177/2331216515617143

Ihlefeld, A., Carlyon, R. P., Kan, A., Churchill, T. H., & Litovsky, R. Y. (2015). Limitations on monaural and binaural temporal processing in bilateral cochlear implant listeners. *Journal of the Association for Research in Otolaryngology*, *16*(5), 641–652. https://doi.org/10.1007/s10162-015-0527-7

Ihlefeld, A., Kan, A., & Litovsky, R. Y. (2014). Across-frequency combination of interaural time difference in bilateral cochlear implant listeners. *Frontiers in Systems Neuroscience*, *8*. https://doi.org/10.3389/fnsys.2014.00022

Joshi, S. N., Dau, T., & Epp, B. (2017). A Model of Electrically Stimulated Auditory Nerve Fiber Responses with Peripheral and Central Sites of Spike Generation. *Journal of the Association for Research in Otolaryngology*, *18*(2), 323–342. https://doi.org/10.1007/s10162-016-0608-2

Kan, A., Jones, H. G., & Litovsky, R. Y. (2015). Effect of multi-electrode configuration on sensitivity to interaural timing differences in bilateral cochlear-implant users. *The Journal of the Acoustical Society of America*, *138*(6), 3826–3833. https://doi.org/10.1121/1.4937754

Kan, A., Stoelb, C., Litovsky, R. Y., & Goupell, M. J. (2013). Effect of mismatched place-of-stimulation on binaural fusion and lateralization in bilateral cochlear-implant users. *The Journal of the Acoustical Society of America*, *134*(4), 2923–2936. https://doi.org/10.1121/1.4820889

Karg, S., Lackner, C., & Hemmert, W. (2013). Temporal interaction in electrical hearing elucidates auditory nerve dynamics in humans. *Hearing Research*, *299*, 10–18. https://doi.org/10.1016/j.heares.2013.01.015

Kerber, S., & Seeber, B. U. (2013). Localization in Reverberation with Cochlear Implants: Predicting Performance from Basic Psychophysical Measures. *Journal of the Association for Research in Otolaryngology*, *14*(3), 379–392. https://doi.org/10.1007/s10162-013-0378-z

Kidd, G., Mason, C. R., Best, V., & Marrone, N. (2010). Stimulus factors influencing spatial release from speech-on-speech masking. *The Journal of the Acoustical Society of America*, *128*(4), 1965–1978. https://doi.org/10.1121/1.3478781

Klingel, M., & Laback, B. (2022). Reweighting of Binaural Localization Cues in Bilateral Cochlear-Implant Listeners. *Journal of the Association for Research in Otolaryngology*, *23*(1), 119–136. https://doi.org/10.1007/s10162-021-00821-3

Kong, Y.-Y., & Carlyon, R. P. (2010). Temporal pitch perception at high rates in cochlear implants. *The Journal of the Acoustical Society of America*, *127*(5), 3114–3123. https://doi.org/10.1121/1.3372713

Kong, Y.-Y., Deeks, J. M., Axon, P. R., & Carlyon, R. P. (2009). Limits of temporal pitch in cochlear implants. *The Journal of the Acoustical Society of America*, *125*(3), 1649–1657. https://doi.org/10.1121/1.3068457

Laback, B., & Majdak, P. (2008). Binaural jitter improves interaural time-difference sensitivity of cochlear implantees at high pulse rates. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(2), 814–817. https://doi.org/10.1073/pnas.0709199105

Laback, B., Egger, K., & Majdak, P. (2015). Perception and coding of interaural time differences with bilateral cochlear implants. *Hearing Research*, *322*, 138–150. https://doi.org/10.1016/j.heares.2014.10.004

Laback, B., Majdak, P., & Baumgartner, W.-D. (2007). Lateralization discrimination of interaural time delays in four-pulse sequences in electric and acoustic hearing. *The Journal of the Acoustical Society of America*, *121*(4), 2182–2191. https://doi.org/10.1121/1.2642280

Laback, B., Pok, S.-M., Baumgartner, W.-D., Deutsch, W. A., & Schmid, K. (2004). Sensitivity to interaural level and envelope time differences of two bilateral cochlear implant listeners using clinical sound processors. *Ear and Hearing*, *25*(5), 488–500. https://doi.org/10.1097/01.aud.0000145124.85517.e8

Laback, B., Zimmermann, I., Majdak, P., Baumgartner, W.-D., & Pok, S.-M. (2011). Effects of envelope shape on interaural envelope delay sensitivity in acoustic and electric hearing. *The Journal of the Acoustical Society of America*, *130*(3), 1515–1529. https://doi.org/10.1121/1.3613704

Landsberger, D. M., Stupak, N., Spitzer, E. R., Entwisle, L., Mahoney, L., Waltzman, S. B., McMenomey, S., Friedmann, D. R., Svirsky, M. A., Shapiro, W., & Roland, J. T. (2022). Stimulating the Cochlear Apex Without Longer Electrodes: Preliminary Results With a New Approach. *Otology & Neurotology*, *43*(5), e578–e581. https://doi.org/10.1097/MAO.0000000000003529

Laneau, J., & Wouters, J. (2004). Multichannel Place Pitch Sensitivity in Cochlear Implant Recipients. *Journal of the Association for Research in Otolaryngology*, *5*(3), 285–294. https://doi.org/10.1007/s10162-004-4049-y

Laneau, J., Wouters, J., & Moonen, M. (2006). Improved music perception with explicit pitch coding in cochlear implants. *Audiology and Neurotology*, *11*(1), 38–52. https://doi.org/10.1159/000088853

Lefebvre, P. P., Müller, J., Mark, G., Schwarze, F., & Hochmair, I. (2025). Rehabilitation of human hearing with a totally implantable cochlear implant: A feasibility study. *Communications Medicine*, *5*(1), 10. https://doi.org/10.1038/s43856-024-00719-0

Lindenbeck, M. J., Frohmann, L., Middlebrooks, J. C., & Laback, B. (2023). *Synergy of rate-pitch and interaural-time-difference cues for voluntary stream segregation in bilateral cochlear-implant listeners* [Presentation]. Conference on Implantable Auditory Prostheses, Lake Tahoe, CA, USA.

Lindenbeck, M. J., & Laback, B. (2025). *Insights into the periodicity encoded by amplitude modulation and short inter-pulse intervals using a phenomenological auditory nerve model* [Presentation]. DAS | DAGA 2025 - 51st Annual Meeting on Acoustics, Copenhagen, Denmark.

Lindenbeck, M. J., Laback, B., Majdak, P., & Srinivasan, S. (2020). Temporal-pitch sensitivity in electric hearing with amplitude modulation and inserted pulses with short inter-pulse intervals. *The Journal of the Acoustical Society of America*, *147*(2), 777–793. https://doi.org/10.1121/10.0000610

Lindenbeck, M. J., Majdak, P., & Laback, B. (2019). Binaural Processing in a New Cochlear-Implant Paradigm Inserting Extra Pulses with Short Inter-Pulse Intervals. *Proceedings of the ICA 2019 and EAA Euroregio : 23rd International Congress on*

Acoustics, *integrating 4th EAA Euroregio 2019 : 9-13 September 2019*, 2202–2209. https://doi.org/10.18154/RWTH-CONV-238847

Lindenbeck, M. J., Majdak, P., & Laback, B. (2024). Effects of Monaural Temporal Electrode Asynchrony and Channel Interactions in Bilateral and Unilateral Cochlear-Implant Stimulation. *Trends in Hearing*, *28*, 23312165241271340. https://doi.org/10.1177/23312165241271340

Lindenbeck, M. J., Majdak, P., & Laback, B. (2025). Effects of Monaural Temporal Electrode Asynchrony on Temporal Pitch Discrimination with Cochlear Implants [manuscript submitted for publication]. *Acoustics Research Institute, Austrian Academy of Sciences*.

Lindenbeck, M. J., Majdak, P., Srinivasan, S., & Laback, B. (2023). Pitch discrimination in electric hearing with inconsistent and consistent amplitude-modulation and inter-pulse rate cues. *The Journal of the Acoustical Society of America*, *153*(6), 3268–3280. https://doi.org/10.1121/10.0019452

Loizou, P. C., Poroy, O., & Dorman, M. (2000). The effect of parametric variations of cochlear implant processors on speech understanding. *The Journal of the Acoustical Society of America*, *108*(2), 790–802. https://doi.org/10.1121/1.429612

Lorenzi, C., Gilbert, G., Carn, H., Garnier, S., & Moore, B. C. J. (2006). Speech perception problems of the hearing impaired reflect inability to use temporal fine structure. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(49), 18866–18869. https://doi.org/10.1073/pnas.0607364103

Macherey, O., & Carlyon, R. P. (2014). Re-examining the upper limit of temporal pitch. *The Journal of the Acoustical Society of America*, *136*(6), 3186–3199. https://doi.org/10.1121/1.4900917

Macherey, O., Carlyon, R. P., van Wieringen, A., Deeks, J. M., & Wouters, J. (2008). Higher Sensitivity of Human Auditory Nerve Fibers to Positive Electrical Currents. *Journal of the Association for Research in Otolaryngology*, *9*(2), 241–251. https://doi.org/10.1007/s10162-008-0112-4

Macpherson, E. A., & Middlebrooks, J. C. (2002). Listener weighting of cues for lateral angle: The duplex theory of sound localization revisited. *The Journal of the Acoustical Society of America*, *111*(5), 2219. https://doi.org/10.1121/1.1471898

Majdak, P., Laback, B., & Baumgartner, W.-D. (2006). Effects of interaural time differences in fine structure and envelope on lateral discrimination in electric hearing. *The Journal of the Acoustical Society of America*, *120*(4), 2190–2201. https://doi.org/10.1121/1.2258390

McKay, C. M., McDermott, H. J., & Clark, G. M. (1994). Pitch percepts associated with amplitude-modulated current pulse trains in cochlear implantees. *The Journal of the Acoustical Society of America*, *96*(5), 2664–2673. https://doi.org/10.1121/1.411377

MED-EL. (2024, June 25). *9 Reasons Why Cochlear Implant Electrodes Matter*. MED-EL Professionals Blog. Retrieved November 24, 2024, from https://blog.medel.pro/surgical/why-cochlear-implant-electrodes-matter/

Middlebrooks, J. C. (2004). Effects of cochlear-implant pulse rate and inter-channel timing on channel interactions and thresholds. *The Journal of the Acoustical Society of America*, *116*(1), 452–468. https://doi.org/10.1121/1.1760795

Middlebrooks, J. C., & Onsan, Z. A. (2012). Stream segregation with high spatial acuity. *The Journal of the Acoustical Society of America*, *132*(6), 3896–3911. https://doi.org/10.1121/1.4764879

Middlebrooks, J. C., & Snyder, R. L. (2010). Selective Electrical Stimulation of the Auditory Nerve Activates a Pathway Specialized for High Temporal Acuity. *The Journal of Neuroscience*, *30*(5), 1937–1946. https://doi.org/10.1523/JNEUROSCI.4949-09.2010

Milczynski, M., Wouters, J., & van Wieringen, A. (2009). Improved fundamental frequency coding in cochlear implant signal processing. *The Journal of the Acoustical Society of America*, *125*(4), 2260–2271. https://doi.org/10.1121/1.3085642

Miller, C. A., Abbas, P. J., & Robinson, B. K. (2001). Response Properties of the Refractory Auditory Nerve Fiber. *Journal of the Association for Research in Otolaryngology*, *2*(3), 216–232. https://doi.org/10.1007/s101620010083

Monaghan, J. J. M., & Seeber, B. U. (2016). A method to enhance the use of interaural time differences for cochlear implants in reverberant environments. *The Journal of the Acoustical Society of America*, *140*(2), 1116–1129. https://doi.org/10.1121/1.4960572

Moore, B. C. J., & Carlyon, R. P. (2005). Perception of pitch by people with cochlear hearing loss and by cochlear implant users. In C. J. Plack, R. R. Fay, A. J. Oxenham, & A. N. Popper (Eds.), *Pitch* (pp. 234–277, Vol. 24). Springer. https://doi.org/10.1007/0-387-28958-5_7

Noel, V. A., & Eddington, D. K. (2013). Sensitivity of bilateral cochlear implant users to fine-structure and envelope interaural time differences. *The Journal of the Acoustical Society of America*, *133*(4), 2314–2328. https://doi.org/10.1121/1.4794372

Oxenham, A. J. (2012). Pitch Perception. *The Journal of Neuroscience*, *32*(39), 13335–13338. https://doi.org/10.1523/jneurosci.3815-12.2012

Oxenham, A. J., Micheyl, C., Keebler, M. V., Loper, A., & Santurette, S. (2011). Pitch perception beyond the traditional existence region of pitch. *Proceedings of the National Academy of Sciences*, *108*(18), 7629–7634. https://doi.org/10.1073/pnas.1015291108

Risi, F. (2019). Considerations and Rationale for Cochlear Implant Electrode Design - Past, Present and Future. *The Journal of International Advanced Otology*, *14*(3), 382–391. https://doi.org/10.5152/iao.2018.6372

Riss, D., Hamzavi, J.-S., Blineder, M., Honeder, C., Ehrenreich, I., Kaider, A., Baumgartner, W.-D., Gstoettner, W., & Arnoldner, C. (2014). FS4, FS4-p, and FSP: A 4-month crossover study of 3 fine structure sound-coding strategies. *Ear and Hearing*, *35*(6), e272–e281. https://doi.org/10.1097/aud.0000000000000063

Riss, D., Hamzavi, J.-S., Selberherr, A., Kaider, A., Blineder, M., Starlinger, V., Gstoettner, W., & Arnoldner, C. (2011). Envelope Versus Fine Structure Speech Coding Strategy: A Crossover Study. *Otology & Neurotology*, *32*(7), 1094–1101. https://doi.org/10.1097/MAO.0b013e31822a97f4

Schoenmaker, E., & van de Par, S. (2016). Intelligibility for Binaural Speech with Discarded Low-SNR Speech Components. In P. van Dijk, D. Başkent, E. Gaudrain, E. de Kleine,

A. Wagner, & C. Lanting (Eds.), *Physiology, Psychoacoustics and Cognition in Normal and Impaired Hearing* (pp. 73–81, Vol. 894). Springer International Publishing. https://doi.org/10.1007/978-3-319-25474-6_9

Sharma, R. K., Smetak, M. R., Patro, A., Lindquist, N. R., Perkins, E. L., Holder, J. T., Haynes, D. S., & Tawfik, K. O. (2022). Speech Recognition Performance Differences Between Precurved and Straight Electrode Arrays From a Single Manufacturer. *Otology & Neurotology*, *43*(10), 1149–1154. https://doi.org/10.1097/MAO.0000000000003703

Smith, Z. M. (2010). *Improved sensitivity to interaural time differences with the FAST coding strategy* [Presentation]. 11th International Conference on Cochlear Implants and Other Implantable Auditory Technologies, Stockholm, Sweden.

Smith, Z. M., Delgutte, B., & Oxenham, A. J. (2002). Chimaeric sounds reveal dichotomies in auditory perception. *Nature*, *416*(6876), 87–90. https://doi.org/10.1038/416087a

Srinivasan, S., Laback, B., Majdak, P., & Arnoldner, C. (2020). Improving interaural time difference sensitivity using short inter-pulse intervals with amplitude-modulated pulse trains in bilateral cochlear implants. *Journal of the Association for Research in Otolaryngology*, *21*, 105–120. https://doi.org/10.1007/s10162-020-00743-6

Srinivasan, S., Laback, B., Majdak, P., & Delgutte, B. (2018). Introducing short interpulse intervals in high-rate pulse trains enhances binaural timing sensitivity in electric hearing. *Journal of the Association for Research in Otolaryngology*, *19*(3), 301–315. https://doi.org/10.1007/s10162-018-0659-7

Stahl, P., Macherey, O., Meunier, S., & Roman, S. (2016). Rate discrimination at low pulse rates in normal-hearing and cochlear implant listeners: Influence of intracochlear stimulation site. *The Journal of the Acoustical Society of America*, *139*(4), 1578–1591. https://doi.org/10.1121/1.4944564

Takanen, M., & Seeber, B. U. (2022). A Phenomenological Model Reproducing Temporal Response Characteristics of an Electrically Stimulated Auditory Nerve Fiber. *Trends in Hearing*, *26*, 233121652211170. https://doi.org/10.1177/23312165221117079

Takanen, M., Strahl, S., & Schwarz, K. (2024). Insights Into Electrophysiological Metrics of Cochlear Health in Cochlear Implant Users Using a Computational Model. *Journal of the Association for Research in Otolaryngology*, *25*(1), 63–78. https://doi.org/10.1007/s10162-023-00924-z

Thakkar, T., Kan, A., Jones, H. G., & Litovsky, R. Y. (2018). Mixed stimulation rates to improve sensitivity of interaural timing differences in bilateral cochlear implant listeners. *The Journal of the Acoustical Society of America*, *143*(3), 1428–1440. https://doi.org/10.1121/1.5026618

Thakkar, T., Kan, A., & Litovsky, R. Y. (2023). Lateralization of interaural time differences with mixed rates of stimulation in bilateral cochlear implant listeners. *The Journal of the Acoustical Society of America*, *153*(3), 1912–1923. https://doi.org/10.1121/10.0017603

Townshend, B., Cotter, N., Van Compernolle, D., & White, R. L. (1987). Pitch perception by cochlear implant subjects. *The Journal of the Acoustical Society of America*, *82*(1), 106–115. https://doi.org/10.1121/1.395554

van Hoesel, R. J. M. (2004). Exploring the Benefits of Bilateral Cochlear Implants. *Audiology and Neurotology*, *9*(4), 234–246. https://doi.org/10.1159/000078393

van Hoesel, R. J. M., Böhm, M., Pesch, J., Vandali, A., Battmer, R. D., & Lenarz, T. (2008). Binaural speech unmasking and localization in noise with bilateral cochlear implants using envelope and fine-timing based strategies. *The Journal of the Acoustical Society of America*, *123*(4), 2249–2263. https://doi.org/10.1121/1.2875229

van Hoesel, R. J. M., Jones, G. L., & Litovsky, R. Y. (2009). Interaural time-delay sensitivity in bilateral cochlear implant users: Effects of pulse rate, modulation rate, and place of stimulation. *Journal of the Association for Research in Otolaryngology*, *10*(4), 557–567. https://doi.org/10.1007/s10162-009-0175-x

van Wieringen, A., Carlyon, R. P., Long, C. J., & Wouters, J. (2003). Pitch of amplitude-modulated irregular-rate stimuli in acoustic and electric hearing. *The Journal of the Acoustical Society of America*, *114*(3), 1516–1528. https://doi.org/10.1121/1.1577551

Wallhagen, M. I. (2010). The Stigma of Hearing Loss. *The Gerontologist*, *50*(1), 66–75. https://doi.org/10.1093/geront/gnp107

Wilson, B. S., Finley, C. C., Lawson, D. T., Wolford, R. D., Eddington, D. K., & Rabinowitz, W. M. (1991). Better speech recognition with cochlear implants. *Nature*, *352*(6332), 236–238. https://doi.org/10.1038/352236a0

Zeng, F.-G. (2002). Temporal pitch in electric hearing. *Hearing Research*, *174*(1–2), 101–106. https://doi.org/10.1016/s0378-5955(02)00644-5

Zeng, F.-G., Nie, K., Liu, S., Stickney, G., Del Rio, E., Kong, Y.-Y., & Chen, H. (2004). On the dichotomy in auditory perception between temporal envelope and fine structure cues (L). *The Journal of the Acoustical Society of America*, *116*(3), 1351–1354. https://doi.org/10.1121/1.1777938

Zeng, F.-G., Nie, K., Stickney, G. S., Kong, Y.-Y., Vongphoe, M., Bhargave, A., Wei, C., & Cao, K. (2005). Speech recognition with amplitude and frequency modulations. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(7), 2293–2298. https://doi.org/10.1073/pnas.0406460102

Zirn, S., Arndt, S., Aschendorff, A., Laszig, R., & Wesarg, T. (2016). Perception of Interaural Phase Differences With Envelope and Fine Structure Coding Strategies in Bilateral Cochlear Implant Users. *Trends in Hearing*, *20*, 1–12. https://doi.org/10.1177/2331216516665608

Zwislocki, J., & Feldman, R. S. (1956). Just Noticeable Differences in Dichotic Phase. *The Journal of the Acoustical Society of America*, *28*(5), 860–864. https://doi.org/10.1121/1.1908495