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 (F_0) in the envelope of a high-rate carrier, were used. Two SIPI-insertion approaches, five modulation depths, two typical speech- F_0 s, and two carrier rates were tested. SIPI pulses were inserted either in every amplitude-modulation period (full-rate SIPI) to support the F_0 cue or in every other amplitude-modulation period (half-rate SIPI) to circumvent a potential rate limitation at higher F_0 s. The results demonstrate that full-rate SIPI pulses improve temporal-pitch sensitivity across F_0 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.

© 2020 Acoustical Society of America. <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 (F_0) is encoded temporally via the envelope repetition rate of auditory-filter outputs from groups of higher (unresolved) harmonics (residue pitch, Schouten, 1940). The F_0 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.

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_0 s 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_0 s and SIPI rates were also tested for comparison with the high-rate conditions and to link the findings to a wider range of literature.

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_0 s, 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_0 s to be discriminated (see Sec. II A 3). The tested nominal F_0 s 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_0 s 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_0 s 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.

Listener	Gender	Age at testing (years)	Etiology	Age at onset of deafness (years)		Ear used for the experiment		
				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	-

(A) No-SIPI Condition:

(B) Full-Rate-SIPI Condition:

(C) Half-Rate-SIPI Condition:

(D) Low-Rate Condition:

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 (F_0). 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 F_0 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 F_0 . Panel (D), low-rate condition: unmodulated pulse trains, the pulse rate corresponds to F_0 , 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 μ s 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 F_0 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 μ s and there was no inter-phase gap. For CI21, the phase duration had to be increased to 33.3 μ s 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,

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_0 s 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_0 s 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 F_0 s.

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_0 s within a trial (cf. Kreft *et al.*, 2010). Note that, due to the definition of the SIPI fraction, the F_0 s were allowed to be only integer sub-multiples of the carrier rate. Thus, the actual average F_0 s 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

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 trials, 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_0 s. 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).

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_0 s from the targeted nominal F_0 s 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_0 s. 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_0 s 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_0 s. 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_0 s 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

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).

Listener	2000-pps carrier rate								1000-pps carrier rate			
	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

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_0 s. Half-rate-SIPI (filled diamonds) performance was broadly similar across F_0 s. 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

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 $F0$ 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 ≥ 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- $F0$ 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 ≥ 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- $F0$ 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

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_0 s 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_0 s [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).

$F0$ 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 μ s 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.

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 $\times$ modulation depth: $F(4,16) = 1.79$, $p = 0.180$, $\eta_p^2 = 0.309$; carrier rate $\times$ stimulus type: $F(2,8) = 1.23$, $p = 0.340$, $\eta_p^2 = 0.236$; carrier rate $\times$ modulation depth $\times$ 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,

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- $F0$ 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. $F0$ s 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 $F0$ s and SIPI rates were tested. A follow-up experiment (experiment 2) tested the low- $F0$ 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 $F0$ s) and 0.3 (high nominal $F0$ 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- $F0$ 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

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 to determine the extent of the difference in loudness perception between the two approaches.

3. SIPI pulses inserted at half rate

For low F_0 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 F_0 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 F_0 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 F_0 , the half-rate-SIPI condition performed significantly worse than the full-rate-SIPI condition, and for the high F_0 , 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 F_0 -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 F_0 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 F_0 . 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 F_0 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 F_0 (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

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 $F0$ s, 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 $F0$ 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 $F0$ s 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

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 μ s 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_0 s. 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_0 s. Thus, there seems to be not much potential for the use of half-rate SIPI pulses.
- (4) For F_0 s 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 μ s.

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.

¹Note 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.

²English 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

- 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 Oberzutz, 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.
- Garofalo, 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.
- Ilhfeld, 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. Neurotol.* **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.

- 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., Blünder, M., Honeder, C., Ehrenreich, I., Kaider, A., Baumgartner, W.-D., Gstöettner, 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 Compernelle, 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.