Research Article

Improving Interaural Time Difference Sensitivity Using Short Inter-pulse Intervals with Amplitude-Modulated Pulse Trains in Bilateral Cochlear Implants

SRIDHAR SRINIVASAN,¹ BERNHARD LABACK,^{1,2} PIOTR MAJDAK,¹ AND CHRISTOPH ARNOLDNER²

¹*Acoustics Research Institute, Austrian Academy of Sciences, Wohllebengasse 12-14, A-1040, Vienna, Austria*

²*HNO Department, Medical University of Vienna, Spitalgasse 23, A-1090, Vienna, Austria*

Received: 2 July 2019; Accepted: 22 January 2020; Online publication: 10 February 2020

ABSTRACT

Interaural time differences (ITDs) at low frequencies are important for sound localization and spatial speech unmasking. These ITD cues are not encoded in commonly used envelope-based stimulation strategies for cochlear implants (CIs) using high pulse rates. However, ITD sensitivity can be improved by adding extra pulses with short inter-pulse intervals (SIPIs) in unmodulated high-rate trains. Here, we investigated whether this improvement also applies to amplitude-modulated (AM) high-rate pulse trains. To this end, we systematically varied the temporal position of SIPI pulses within the envelope cycle (SIPI phase), the fundamental frequency (F0) of AM (125 Hz and 250 Hz), and AM depth (from 0.1 to 0.9). Stimuli were presented at an interaurally place-matched electrode pair at a reference pulse rate of 1000 pulses/s. Participants performed an ITD-based left/right discrimination task. SIPI insertion resulted in improved ITD sensitivity throughout the range of modulation depths and for both male and female F0s. The improvements were largest for insertion at and around the envelope peak. These results are promising for conveying salient ITD cues at high pulse rates commonly used to encode speech information.

Keywords: bilateral cochlear implant, binaural timing cues, ITD sensitivity, amplitude modulation,

high-rate stimulation, short inter-pulse intervals

INTRODUCTION

Current cochlear implants (CIs) are quite successful in restoring speech understanding in quiet in the deaf or hard-of-hearing. They are, however, only moderately successful in providing spatial hearing and largely fail in providing listeners with the ability to understand a speaker in a noisy environment, e.g., with interfering speaker(s). While normal hearing (NH) listeners use binaural cues, besides monaural cues, to segregate speech sources, CIs fail to sufficiently provide such cues. Here, we study an approach to provide salient temporal cues for spatial hearing with modulated high-rate pulse trains that are commonly used to encode speech.

Azimuthal sound localization is based on binaural cues, the so-called interaural time differences (ITDs) and interaural level differences (ILDs). The two cues are to some extent complementary, with ITDs contributing mainly for sounds containing low frequencies and ILDs only contributing for sounds containing high frequencies (Macpherson and Middlebrooks 2002; Strutt 1876). Although ITDs arise both in the carrier signal (or fine structure) and the envelope of modulated signals, fine-structure ITD at lower frequencies has been shown to contribute most to azimuthal localization (Macpherson and Middlebrooks 2002; Wightman and Kistler 1992). Envelope ITD can, however, also provide salient azimuthal localization cues if the envelope is sufficiently sharp in terms of its modulation depth, slope,

The authors are ordered according to their contribution to this study.

Correspondence to: Bernhard Laback · Acoustics Research Institute · Austrian Academy of Sciences · Wohllebengasse 12-14, A-1040, Vienna, Austria. email: bernhard.laback@oeaw.ac.at

and duty cycle (Bernstein and Trahiotis 2002; Klein-Hennig et al. 2011; Laback et al. 2011). ITDs are pivotal for perceptual segregation of horizontally separated sound sources (e.g., Middlebrooks and Onsan 2012) and, consequently, for obtaining spatial release from speech-in-speech masking (e.g., Ihlefeld and Litovsky 2012; Kidd et al. 2010).

Cochlear implant listeners are often sensitive to ITD presented directly to the implants under precise stimulus control using a research interface. Best sensitivity is obtained for unmodulated low-rate pulse trains or high-rate pulse trains with low-rate AM (van Hoesel et al. 2009; for a review, see Laback et al. 2015). Similar to normal hearing, envelope ITD sensitivity for AM pulse trains depends on the modulation shape (Laback et al. 2011). ITD sensitivity generally deteriorates with increasing rate beyond about 200 pulses per second (pps), the rate referring either to the carrier pulses in case of no AM or to the modulator in case of high carrier pulse rates (van Hoesel et al. 2009; Laback et al. 2007; Noel and Eddington 2013; Srinivasan et al. 2018). This so-called ITD rate limitation shares some similarities with a rate limitation for envelope ITD in normal hearing (e.g., Bernstein and Trahiotis 2002). Notably, fine-structure ITD sensitivity in normal hearing is not affected by such a rate limitation (Brughera et al. 2013; Klumpp and Eady 1956), further contributing to an apparent deterioration in CI compared to NH listeners (see Laback et al. 2015).

Widespread envelope-based CI processing strategies, like the continuous interleaving sampling (CIS; Wilson et al. 1991), generally discard fine-structure ITD but transmit envelope ITD cues. This is because these strategies use high-rate carrier signals that are not time-locked to the acoustic input signal, and, even if they would be time-locked, their rate is much too high to convey any useful carrier ITD cues. Envelope ITD sensitivity of bilateral CI listeners using CIS is generally weak (Grantham et al. 2008; Laback et al. 2004); thus, it is not surprising that sound localization performance under such conditions relies heavily on ILD while envelope ITD contributes little (Grantham et al. 2007; Seeber and Fastl 2008). Some more recent approaches attempt to better transmit ITD cues by including low-rate channels where pulses are time-locked to the acoustic input signals (FSP; Hochmair et al. 2006, PDT; van Hoesel et al. 2008, FS4; Riss et al. 2014, FAST; Smith 2010, Thakkar et al. 2018, Williges et al. 2018). However, to date, there are no published clinical or laboratory stimulation strategies showing improvements in sound localization compared to high-rate strategies or showing evidence for spatial release from masking for speech. The goal to provide salient speech and ITD cues at the same electrode actually raises a dilemma: encoding the speech

envelope requires pulse rates of at least a few hundred pps (e.g., Arora et al. 2009; Loizou et al. 2000), while ITD sensitivity is already largely degraded by the ITD rate limitation at such rates (see also Churchill et al. 2014). Note that encoding ITD and speech information on separate electrodes is no satisfactory solution because it further reduces the number of available electrode for speech and ITD coding.

An approach to increase ITD sensitivity at high pulse rates was proposed in Laback and Majdak (2008), by introducing random but interaurally coordinated jitter in the pulse timing. A follow-up physiological study by Hancock et al. (2012) showed that the increased sensitivity resulted from highly synchronized neuronal firing triggered by randomly occurring short inter-pulse intervals (SIPIs). This idea was further investigated in psychophysical (Srinivasan et al. 2018) and physiological (Buechel et al. 2018) studies, testing the effect on ITD sensitivity of periodically adding extra pulses to unmodulated high-rate pulse trains (1000 pps), thus creating SIPIs. Systematic variation of the rate of the extra pulses (i.e., the SIPI rate) and the temporal position of the extra pulses relative to the regular carrier pulses (i.e., the SIPI fraction) revealed SIPI conditions completely restoring ITD sensitivity to the level obtained with unmodulated low-rate pulse trains (Srinivasan et al. 2018).

Because these two SIPI studies employed unmodulated pulse trains, they leave unanswered the question if SIPI pulses are also beneficial for more realistic stimuli involving AM, which may already provide ITD cues in the envelope. Hence, the primary question in this study is whether introduction of SIPIs in speech-like AM high-rate pulse trains provides benefits in addition to the ITD cues provided by the AM. Given that the best envelope ITD sensitivity at approximately 100–200 Hz (Noel and Eddington 2013) overlaps well with the fundamental frequency (F0) range of human voiced speech, we employed F0s corresponding to mean male and female speech (125 and 250 Hz, respectively). This allowed us to test the conditional hypothesis that if the higher F0 is affected by a ITD rate limitation (which was not clear a priori), the SIPI-based benefit is larger at that higher F0 due to complete, F0-independent restoration of highly stimulus-synchronized AN spikes by the SIPI pulses. To test that hypothesis, we fixed the SIPI rate across F0s.

An important parameter is the phase of the AM cycle at which SIPI pulses should be inserted in order to provide maximum benefit. There are different indications that the peak phase may be optimal for placing SIPI pulses. First, loudness of AM pulse trains in electric hearing suggests a higher perceptual weight of the modulation peak (McKay and Henshall

2010). Second, and related to the first point, ITD sensitivity in electric hearing increases with pulse amplitude (Egger et al. 2017). Third, perceptual weight of ITD information appears to be highest at or around the modulation peak, both in NH listeners presented with high-frequency stimuli with low-rate AM and CI listeners presented with low-rate electric pulse trains (Hu et al. 2017). Because the outcomes might differ for our high-rate AM pulse trains, we systematically varied the SIPI phase across the positive modulation cycle.

Finally, the modulation depth of real-world speech sounds is often largely reduced, due to CI-specific signal processing, to reverberation or to background noise. To provide a more complete picture, we therefore tested all conditions described so far as a function of modulation depth. It was hypothesized that insertion of SIPI pulses enhances ITD sensitivity particularly at low modulation depths where the envelope ITD cues are weak.

METHODS

For each listener, an interaurally place-matched electrode pair was determined, for which binaural comfortable levels were estimated (more details in the “Binaural Electrode Pair, Electric Dynamic Range, and Binaural Levels” section). Pitch-matched electrodes were used because they were more likely to elicit better ITD sensitivity than non-matched elec-

trodes (Kan et al. 2015; Poon et al. 2009). The binaural comfortable levels aimed at producing an auditory image that was perceived as centered. These levels were used as reference in the loudness-matching procedures for the experimental stimuli. For the main test on ITD sensitivity, a range of optimal parameters was determined in an ITD pretest, in order to avoid floor and ceiling effects across all experimental conditions for each individual listener in the main ITD test.

Listeners and Apparatus

Seven (six postlingually deafened and one prelingually deafened) listeners bilaterally implanted with 12-electrode CIs (Med-El Inc., Austria) participated in the experiment. The inclusion criteria for participation in the study were good speech perception in quiet and a better-than-chance ITD sensitivity for unmodulated low-rate pulse trains and unmodulated high-rate pulse trains with SIPI pulses (more details in the “ITD Sensitivity” section). Individual listener data are shown in Table 1. All listeners were paid an hourly wage for their participation.

Stimuli were generated on a personal computer and presented to the CIs via a research interface (RIB2, Institute of Ion Physics and Applied Physics, Leopold-Franzens University of Innsbruck, Austria). The RIB2 allows direct and interaurally coordinated stimulation of two CIs. All procedures presented here were approved by the ethics committee of the Medical University of Vienna (vote #2155/2013).

TABLE 1

Participant information regarding age of onset of deafness, implantation of the CI, and electrode pair used for testing. Age of onset of deafness refers to the earlier of the two ears and may also refer to the age at the approximate onset of profound hearing loss when no specific date of onset of deafness could be defined (e.g., in the case of progressive hearing loss)

Listener	Implants (L and R)	Age at testing (years)	Etiology	Age at onset of deafness (years)	Age at implantation (years)		Duration of bilateral stimulation (years)	Electrode pair	
					L	R		L	R
CI1	C40+	31	Meningitis	14	14	14	17	11	10
CI8	C40+	52	Osteogenesis imperfecta	27	41	39	11	10	9
CI12	C40+	49	Vestibular aqueduct syndrome; progressive	29	35	33	14	10	10
CI17	Synchrony	71	Idiopathic audiogram; deterioration	40	67	58	2	8	8
CI24	C40+	53	Progressive	39	41	43	10	8	9
CI88	Pulsar (L); Sonata (R)	65	Meningitis	40	55	57	8	9	7
CI100*	Pulsar (L); C40+ (R)	20		0	8	3	12	9	9
Mean		48.7		27	37.3	35.3	10.6	9.3	8.9

*This listener was excluded from further testing because he/she did not show any sensitivity with high-rate stimuli with SIPI pulses, as one of our selection criteria (see the “Screening” section)

Stimuli and Conditions

Our main stimuli, i.e., pseudo-syllables, aimed to mimic the output of typical CI processor driven with speech syllable inputs, while still being based on a parametric stimulus representation allowing us to systematically vary parameters. To this end, our pseudo-syllables were constructed by amplitude modulating 600-ms trains of biphasic electric pulses. Each phase of a pulse had a duration of 26.7 μ s, there was no between-phase gap, and the carrier pulse rate was of 1000 pps. Linear onset and offset ramps of 150 ms each were applied to minimize transient ITD cues at the onset and offset. The condition using these unmodulated trains is referred to as the *unmodulated reference*. The steady-state portion of the pulse trains was amplitude-modulated by a full-wave rectified sinusoid, roughly resembling the shape of CIS-like processed voiced speech (Dorman and Wilson 2004; Wilson 2006). The stimulus modulation rate after full-wave rectification defines the F0 encoded in the envelope. Conditions using the AM trains are referred to as the *AM reference*, (see Fig. 1a). Further, SIPI pulses were inserted to the steady-state part at the rate described by the SIPI rate. Unmodulated trains with SIPI pulses are referred to as the *SIPI reference* (cf., Srinivasan et al. 2018). In all other *SIPI conditions*, extra pulses were inserted periodically in the AM reference trains with SIPI pulses (see Fig. 1b) parameterized by the SIPI fraction (the time interval between the extra pulse and the preceding pulse, relative to the carrier inter-pulse time interval, in

percent) and the SIPI phase (the temporal position of a SIPI pulse relative to the AM cycle, in degrees) for F0s of 125 Hz (see Fig. 1c) and 250 Hz (see Fig. 1d).

For the choice of F0, we aimed to represent typical fundamental frequencies of either male or female speech. To this end, we analyzed the output of CIS-processed speech from German (Kiel Corpus, Kohler 1996) and English corpora (TIMIT Corpus, Garofolo et al. 1993) and obtained F0 peaks at 210 Hz and 120 Hz. This corresponds to the average of 234 ± 18 and 133 ± 12 Hz of French speakers (Pépiot 2014). Thus, we chose F0s of 125 and 250 Hz. These particular F0s also comply with an integer ratio of carrier rate and F0, required to insert SIPI pulses in each modulation cycle at exactly the same modulation phase.

For the choice of the MD, we aimed to consider realistic conditions. Our CIS-processed speech analysis showed MDs in the range from 0.1 to 0.4 with the median of 0.25. However, ITD sensitivity has been shown to be influenced not only by the modulation depth, but also by the attack slope and the waveform-off time (Klein-Hennig et al. 2011; Laback et al. 2011). In order to include a wider range of effective modulation cues, we chose MDs from 0.1 to 0.9, varied in steps of 0.2. Note that we defined the MD relative to the individual listener's electric dynamic range.

The aim for the SIPI was to maximally enhance the ITD sensitivity for modulated high-rate pulse

FIG. 1. Color online: schematic representation of stimuli shown monaurally in panels **a–d** and binaurally in panel **e**. **a** AM reference high-rate pulse train with F0 of 125 Hz. **b** High-rate pulse train with F0 of 125 Hz with inserted SIPI pulses in every $1/\text{SIPI rate}$ interval; inset shows SIPI fraction. **c** SIPI phase conditions for F0 of 125 Hz. **d**

SIPI phase conditions for F0 of 250 Hz. **e** Binaural illustration of pulse train with SIPI at 90° (from panel **c**), with ITD applied to the entire pulse train. SIPI, short inter-pulse interval; A, amplitude modulation. For more information on the SIPI phase, refer to the text

trains. As for the SIPI rate and SIPI fraction, our choice was based on Srinivasan et al. (2018), who, using unmodulated high-rate pulse trains, showed most improvement with SIPI rates between 50 and 100 pps combined with SIPI fractions between 6 and 10 %. Thus, our SIPI fraction was set to 10 %. Given the requirement that the SIPI rate should be an integer sub-multiple of the modulation rate and the declining efficiency of SIPIs with rates greater than 100 pps, we chose the SIPI rate of 62.5 pps, irrespective of the F0.

For the SIPI phase, the optimum value was not a priori clear. Hu et al. (2017) showed CI listeners being more sensitive to ITD cues presented at the peak of the modulation cycle as compared to its onset. Thus, SIPI pulses added at the modulation peaks seemed to be most promising. Hence, in order to obtain a more complete picture across the modulation cycle, we varied SIPI phase as an independent experimental parameter ranging from 0° (at the off-peak phase of the modulation cycle), over 90° (at the modulation peak) to 180° (being identical with 0°). Note that in contrast to the stimuli of Hu et al. (2017), our stimuli had whole-waveform ITD, thus containing ITD throughout the modulation cycle, although most likely only the SIPI pulse pairs triggered ITD cues. In order to save experiment time, not all combinations of MD and SIPI phase were tested. For the MDs of 0.3 and 0.9, SIPI phases from 0° to 180° were tested in steps of 22.5°, with the assumption that SIPIs would be more efficient in the higher-level region of the modulation period (Hu et al. 2017). For the other MDs tested (0.1, 0.5, and 0.7), a reduced set of SIPI phases was tested. Tables 2 and 3 show all combinations of SIPI phases and MDs tested for the F0 of 125 and 250 Hz, respectively. The amplitude of a SIPI pulse was determined by the instantaneous AM envelope at the time of the SIPI pulse. Note that based on

TABLE 2

Experimental and reference conditions tested for F0s of 125 Hz. Refer to Figs. 1 and 2 and text for explanation of the parameters mentioned

MD SIPI phase (°)	0	0.1	0.3	0.5	0.7	0.9
0			x			x
22.5			x			x
45			x			x
67.5		x	x	x	x	x
90		x	x	x	x	x
112.5		x	x	x	x	x
135			x			x
157.5			x			x
AM reference	x	x	x	x	x	x
SIPI reference	x					

TABLE 3

Experimental and reference conditions tested for F0 of 250 Hz. Refer to Figs. 1 and 4 and text for explanation of the parameters mentioned

MD SIPI phase (°)	0	0.1	0.3	0.5	0.7	0.9
0			x			x
45		x	x	x	x	x
90		x	x	x	x	x
135		x	x	x	x	x
AM reference	x	x	x	x	x	x
SIPI reference	x					

statistical analysis of results (see the “F0 of 125 Hz” section), SIPI phases of 67.5°, 90°, and 112.5° were later grouped into *peak* SIPI phase conditions, whereas other conditions were grouped into off-peak phase SIPI conditions.

Binaural Electrode Pair, Electric Dynamic Range, and Binaural Levels

In order to determine the dynamic range, comfortable binaural levels, and the binaural electrode pair, we performed pretests similar to those from Majdak et al. (2006) and Srinivasan et al. (2018).

The binaural electrode pair was determined using an unmodulated 300-ms periodic pulse trains at a rate of 1515 pps. This high rate was used in order to reduce the potentially confounding effect of rate pitch at lower rates. In a first step, the electric dynamic range in current units (mapped linearly in μA), defined by the threshold (THR) and maximum comfortable level (MCL), as well as an intermediate level that was judged as a comfortable level (CL) for long periods of time was determined for each electrode in both ears using a graphical loudness scale with a verbal category “middle loud” at the center of the scale. Second, a magnitude estimation procedure was used to estimate the perceived pitch across the electrodes at both ears. Stimuli were presented monaurally in random order at either ear and at each of the available electrodes. Listeners were instructed to assign numbers without any restrictions, according to the perceived pitch of each stimulus. Eight or nine candidate interaural electrode pairs were then chosen for a subsequent pitch-ranking procedure. Third, listeners compared members of each candidate interaural electrode pair to indicate which electrode elicited a higher pitch percept. A two-interval, two-alternative forced-choice procedure was used. The pitch-matched pair was selected from the basal region of the cochlea by requiring the pitch-ranking score for that pair to be at chance level. For more details on the pitch-matching paradigm, see

Majdak et al. (2006). This pretest resulted in electrode pairs shown in Table 1.

Then, the monaural dynamic range (described by the MCL and THR) and CLs were determined at the chosen electrode pair for pulse trains having the rate used in the main experiment, i.e., 1000 pps. For the unmodulated trains, we used the same psychoacoustic procedure as for the 1515-pps pulse trains used in the preceding pitch-matching procedure, yielding THR_U , MCL_U , and CL_U . For the modulated trains, the CLs were calculated for each MD separately, yielding $CL_{MD} = CL_U + MD \cdot (CL_U - THR_U)$. In cases where a CL_{MD} exceeded the MCL_U , separate MCL_{MD} were again psychoacoustically determined.

Finally, the binaural levels of the stimuli were established. In order to save experiment time, this was done for the AM references with MD of 0.5 only (the levels of all other stimuli were adjusted in the subsequent loudness-balancing procedure). To this end, AM trains with an MD of 0.5 were presented binaurally, while instructing the listener to attend and respond to the overall loudness by indicating on the visual scale at the perceived loudness. Beginning with 80 % of the monaural $CL_{MD=0.5}$, levels were varied simultaneously in proportionally equal steps at the two ears to finally arrive at the most comfortable binaural CL. We checked if the auditory image was perceived as centered, using a visual horizontal bar with an indication of the midline. If that was not the case, level adjustments were made until the listeners reported a centered auditory image at a CL. These final adjusted binaural CLs were used as the reference in the following loudness-matching procedure.

Loudness-Matching Across Conditions and Centralization

Increasing the pulse rate of electric pulse trains is known to increase the perceived loudness (McKay and McDermott 1998; Shannon 1985). The effect of adding SIPI pulses on loudness, however, is not easily predictable from existing loudness models (McKay and McDermott 1998). In order to avoid confounding effects of loudness on our ITD experiments (e.g., Egger et al. 2016, 2017), we performed a formal loudness-matching procedure for all experimental conditions. An adaptive double staircase loudness-balancing procedure (Jesteadt 1980) was used with the AM reference with MD = 0.5 serving as a *loudness reference*.

The double staircase procedure consisted of a lower and an upper staircases, the starting levels of which were 20 % below and 20 % above the level of the loudness reference, respectively. The levels are described as percentage of the dynamic range defined by MCL_U and THR_U for the particular electrodes on

the left and right ears. The loudness reference and the comparison stimuli were presented in random order in the two intervals of a trial with a silent interval of 200 ms. The listeners reported which of the two stimuli was perceived as louder by pressing the corresponding button. A 3-down 1-up rule was followed for adaptive level adjustment at each step, which converged at the 79 % point of the psychometric function (Levitt 1971). The initial step size was 10 %, and that step size was reduced by a factor of 0.7 after each reversal until it reached the minimum of 2 %. The procedure ended after 12 reversals. The last six reversals from the lower and the upper staircases, converging at the 21 % and 79 % points of the psychometric function, respectively, were averaged to estimate the loudness-matched level. At least two runs were performed to obtain the final loudness-matched stimulation levels.

Once these final loudness-matched stimulation levels were obtained, they were presented to determine if listeners perceived the binaural auditory image as centered, using a visual horizontal bar with an indication of the midline. If that was not the case, small level adjustments were made until the listeners reported a centered auditory image. These adjusted binaural levels were used for the different experimental conditions in the ITD experiments.

ITD Sensitivity

First, in a screening test, we determined the pool of our listeners being sensitive to ITD, particularly for high-rate pulse trains with SIPI pulses. Then, in a pretest, we tested a wide range of ITDs for different conditions to identify a listener-specific ITD encompassing the individual's sensitivity. Finally, in the main test, we determined ITD sensitivity for a given experimental condition based on that listener-specific ITD in terms of the signal detection theory-based sensitivity index d' (Green and Swets 1966). With this setting, we reduced experiment duration while attempting to avoid floor and ceiling effects.

Procedure. A constant-stimuli paradigm was used to measure left/right discrimination performance. A trial consisted of two intervals separated by a 250-ms silent interval. The first interval contained a pulse train stimulus with zero ITD evoking a centered auditory image. The second interval contained the same stimulus with a non-zero ITD. The listeners had to indicate to which side (left or right) the second stimulus was perceived compared with the first stimulus by pressing the corresponding button. The ITD of the target stimulus was applied with equal a priori probability either to the left or to the right side. Visual feedback on the correctness of the response was presented after each trial. A total of 100 trials

were presented per condition, 50 with the target stimulus to the left.

Screening. First, we evaluated whether the listeners were sensitive to ITD cues for (a) unmodulated pulse trains with a low rate, i.e., 100 pps, and (b) unmodulated pulse train with a high rate, i.e., 1000 pps, and SIPI pulses (at a rate of 62.5 pps inserted at a 10 % SIPI fraction). Combined with a non-zero ITD, both conditions were expected to provide salient ITD cues (Srinivasan et al. 2018). From the seven listeners tested, one listener (CI100) had to be excluded from subsequent testing because of not showing any improvements in ITD sensitivity to high-rate pulse trains with SIPI, even though showing good sensitivity to ITD cues in low-rate pulse trains (100 pps). Eventually six listeners completed the study with F0 of 125 Hz, but because of time constraints and availability, only four from that set of listeners further completed the study with the F0 of 250 Hz.

Pretest. A selection of experimental conditions of the main experiment was tested to find a listener-specific ITD encompassing the performance range from chance, i.e., 50 to 100 %, across a variety of conditions. The tested stimuli were peak SIPI conditions with two MDs, namely of 0.3 and 0.9, aimed at triggering the lower and upper bounds of the performance range, respectively. The MD of 0.3 (instead of 0.1) was chosen in order to better capture potential improvements at MDs occurring more frequently in human speech. The stimuli were presented at the loudness-matched amplitudes determined in the loudness-matching pretest. The tested ITDs were chosen from the set of 100, 200, 400, 800, 1200, and 1600 μ s based on the individual performance from the screening.

As a result, the ITDs for the main test were chosen extrapolating from the pretest results to other experiment conditions such that the performance across all experiment conditions would span the whole range from slightly above chance level to near maximum, reducing potential floor and ceiling effects. The selected ITDs are shown in Table 4. For the F0 of 125 Hz, they range from 250 to 1000 μ s (with a mean $\pm$ standard deviation of 599 ± 332 μ s). For the F0 of

250 Hz, they range from 200 to 800 μ s with (470 ± 289 μ s). Note that six listeners performed the experiment at the F0 of 125 Hz, but only a subset of four listeners were tested for the F0 of 250 Hz. For them, the chosen ITDs were very similar across the two F0s (474 ± 241 μ s and 470 ± 289 μ s, for F0s of 125 and 250 Hz, respectively). This reflects similar baseline sensitivity at the two F0s for these four listeners.

Main Experiment. Based on the main ITD pretest, a listener-specific ITD was tested for all experimental conditions grouped within one F0 block. The two F0 frequencies, 125 Hz and 250 Hz, were tested on different days with the F0 of 125 Hz first. Each F0 block consisted of (1) the unmodulated reference, (2) the SIPI reference, (3) the AM references at all MDs, and (4) the SIPI conditions at all considered combinations of MD and SIPI phases. The unmodulated reference and SIPI reference served as control conditions for comparison across F0 blocks. The full set of conditions are shown in Tables 2 and 3 for F0s of 125 Hz (32 conditions) and 250 Hz (24 conditions), respectively.

Within each F0 block, all hundred trials from all conditions were pooled into a single set, and the order of trials was randomized. The set was then divided into blocks of approximately 300 trials each, yielding 11 and 8 blocks for the F0s of 125 and 250 Hz, respectively. Each block lasted approximately 20 min and the listeners took breaks between blocks ad libitum. All testing was done over 3 days.

RESULTS

In general, our approach to pre-select ITD values based on a pretest in order to minimize ceiling or floor effects was successful: the performances across different conditions were well distributed across the range from close-to-chance level up to high levels, with mostly sufficient room at the ceiling and the floor. The percentages of correct left/right discrimination for each experimental condition were converted to d' scores in order to account for potential response bias (Klein 2001). First, we qualitatively describe the results, separately for the F0s of 125 and 250 Hz. Then, we provide an extensive statistical analysis of all effects under investigation.

F0 of 125 Hz

Figure 2 shows individual listeners' d' scores as a function of SIPI phase (filled symbols), with the AM reference conditions (open symbols) on the left side of each panel. Each panel shows a different MD. The average results across listeners are shown with thick

TABLE 4

Tested ITDs across F0s of 125 Hz and 250 Hz

F0 Listener ID	125 Hz ITD in μ s	250 Hz ITD in μ s
CI1	400	400
CI8	650	
CI12	250	200
CI17	800	800
CI24	400	400
CI88	1000

FIG. 2. Color online: for an F_0 of 125 Hz, panels show individual (colored/gray diamonds) and average (black diamonds) d' scores plotted against short inter-pulse interval (SIPI) phase for a certain modulation depth (MD) indicated at bottom of each panel. Error bars show standard errors of the means. AM reference conditions (open

symbols) for that MD are plotted in the left portion of each panel, and as a range in the right. MD, modulation depth; SIPI, short inter-pulse interval; AM, amplitude modulation. For more information on labels MD and SIPI phase, refer to Table 2

diamonds with the error bars indicating standard errors. The standard error ranges of the AM references are also shown across the SIPI phases as blue areas for ease of comparison to the SIPI conditions. Despite listener-specific differences in overall performance, the general pattern across conditions appears to be consistent. Insertion of SIPI pulses around the peak of the AM cycle (SIPI phases of 67.5°, 90°, and 112.5°) enhanced ITD sensitivity. For the off-peak phases (0°, 22.5°, 45°, 135°, and 157.5°), the enhancement decreased towards zero, i.e., the performance was within the range observed for the AM references. Figure 3 (left panel) shows the averages (symbols) and

standard errors (bars) across MDs, demonstrating the global effect of SIPI phase on ITD sensitivity. Note the conditions with narrower error bars for which only data for MDs of 0.3 and 0.9 are available. On a more detailed level, there seems to be an asymmetry of the three phase conditions at the peak (67.5, 90, and 112.5°), with lower performance on the decaying compared to the rising flank.

Figure 3 (right panel) shows the same data but as a function of MD, contrasting performance between the conditions peak SIPI phase (filled diamonds) and off-peak SIPI phase (filled triangle), as well as the AM references (open diamonds). For further comparison,

FIG. 3. Color online: (left panel) for F_0 of 125 Hz, average of all participants across all MDs for the different phases. AM reference conditions (open symbols) are plotted in the left portion of the graph. Error bars indicate standard errors, with conditions underlying a larger number of MDs having thicker bars. Right panel: in the right portion for F_0 of 125 Hz, group average for each participant (colored/gray symbols) and overall average (black symbols) d' scores are plotted against MD (with error bars indicating standard errors). AM reference (open diamonds), average peak SIPI (filled diamonds),

and average off-peak SIPI (inverted filled triangles) d' scores are plotted for each modulation depth. In the left portion, unmodulated no SIPI (open symbols) and SIPI (filled symbols) conditions are plotted. Error bars show standard errors of the means. MD, modulation depth; SIPI, short inter-pulse interval; AM, amplitude modulation; unmod., unmodulated; peak SIPI, SIPI pulse at peak phases; off-peak SIPI, SIPI pulse at off-peak phases; NoSIPI, no SIPI pulses

performances obtained for the unmodulated references are plotted on the left side of the abscissa.

Several observations can be made. First, performance obtained for the AM references increased monotonically with increasing MD. Second, introducing SIPI pulses at the peak SIPI phases improved performance obtained with the F0-based envelope ITD cue alone by about a constant amount across all MDs tested. This amount of improvement appears to be similar to that observed for the unmodulated stimuli. Third, introducing SIPI pulses at off-peak SIPI phases did not change the performance, neither in terms of improvement nor in terms of degradation.

F0 of 250 Hz

Figures 4 and 5 show the results for the F0 of 250 Hz, analogous to Figs. 2 and 3. The overall pattern of results across conditions is similar to that obtained for the F0 of 125 Hz. When compared to the results for F0 of 125 Hz, for small MDs, the peak SIPI effect seems to be larger. The unmodulated conditions (left side of Fig. 5, right panel) show a similarly enhanced SIPI effect. A potential explanation in terms of a training effect is addressed in the “Discussion” section below.

Similar to the F0 of 125 Hz, the off-peak condition shows similar performance to that of the AM reference condition for comparable points (at MDs of 0.3 and 0.9). Finally, the performance in the off-peak condition at MD = 0.1 differs from other MDs in that it approaches that of the peak SIPI condition. This can be understood considering the envelope of this

condition (and, thus, the amplitude of off-peak pulses) approaching that of an unmodulated pulse train (see the “Discussion” section).

Statistical Analysis

The statistical significance of the effects described above was assessed using repeated measures (RM) analysis of variance (ANOVA) and subsequent post hoc tests including the Bonferroni correction for multiple comparisons (multcompare, MATLAB, Mathworks). The Kolmogorov-Smirnov test and the Shapiro-Wilk test were performed on the residuals from the RM-ANOVAs to ensure that they all met the normality criterion (SPSS ver. 26, IBM). The Levene’s test of homogeneity of variance was performed (SPSS ver. 26, IBM), fulfilling the criterion in all cases except for one. In that case, we report a follow-up RM-ANOVA that fulfilled the criterion. All statistical analyses were performed with a significance criterion of 0.05. Results for the parameter combinations shown in Tables 2 and 3 were included in the statistical analyses because only these were available for all participants. Some extra combinations of SIPI phase and AM depth tested for some participants were not included.

F0 of 125 Hz. Starting with the F0 of 125 Hz, a two-way RM-ANOVA was conducted on the factors SIPI phase (all phases) and MD (MDs of 0.3 and 0.9 because for these all phases were tested). Both main effects, SIPI phase ($F_{7,75} = 10.13$, $p < 0.001$) and MD ($F_{1,75} = 36.57$, $p < 0.001$), were significant. However, the interaction between SIPI phase and MD was not significant ($F_{7,75} = 1.53$, $p = 0.174$). Post hoc comparison of the

FIG. 4. Color online: same as in Fig. 2, but for an F0 of 250 Hz. MD, modulation depth; SIPI, short inter-pulse interval; AM, amplitude modulation. For more information on labels MD and SIPI phase, refer to Table 3

FIG. 5. Color online: same as in Fig. 3, but for an F0 of 250 Hz. MD, modulation depth; SIPI, short inter-pulse interval; AM, amplitude modulation; unmod., unmodulated; peak SIPI, SIPI pulse at peak phases; off-peak SIPI, SIPI pulse at off-peak phases; NoSIPI, no SIPI pulses

factor SIPI phase showed a significant difference between the phases at the peak of the envelope and the phases in the trough. Based on these comparisons, the earlier grouping of phases as peak phases (67.5° , 90° , and 112.5°) and as off-peak phases (0° , 22.5° , 45° , 135° , 157.5°) is supported.

An additional two-way RM-ANOVA was conducted on the three phase conditions classified as peak phase, because only these were tested across all MDs, using the factors Phase (67.5° , 90° , and 112.5°) and MD (all values tested). Results showed significant main effects of Phase ($F_{2,70} = 5.8$, $p = 0.005$) and MD ($F_{4,70} = 22.91$, $p < 0.001$), but no significant interaction ($F_{8,70} = 1.24$, $p = 0.288$). Post hoc tests showed significantly lower performance for the 112.5° compared to both the 67.5° and 90° conditions. Note that these differences, albeit significant, are minor compared to the difference between the peak and off-peak phases.

Then, the conditions AM reference and peak phase were compared across all MDs using a two-way RM-ANOVA with the factors Condition (levels: peak phase and AM reference) and MD (all tested levels, including the unmodulated condition). Both main effects, Condition ($F_{1,115} = 108.57$, $p < 0.001$) and MD ($F_{5,115} = 17.45$, $p < 0.001$) were significant, while the interaction between these factors was not significant ($F_{5,115} = 0.61$, $p = 0.693$). Post hoc tests showed that the difference between AM reference and peak phase conditions was significant at all MDs, except in the unmodulated conditions. Further, post hoc comparisons for the AM reference condition showed significantly higher performance for the MD of 0.9 compared to the MD of 0.

While in the previous analysis we did not include the off-peak phase conditions because it was tested only for a subset of MDs and, thus, would have restricted the evaluation of the peak phase SIPI effect across MDs, the potential of off-peak SIPI

pulses still needs to be clarified. To this end, a two-way RM-ANOVA with the factors Condition (AM reference, peak phase, off-peak phase) and MD (0.3 and 0.9) was performed. However, this ANOVA was reconsidered because it did not satisfy the homogeneity of variance assumption. Given the lack of any apparent interaction between Condition and MD from a visual inspection of the available data, the data were averaged across the two modulation depths and a one-way RM-ANOVA with the factor Condition (AM reference, peak phase, off-peak phase) was performed. This ANOVA satisfied the homogeneity of variance criterion. The main effect of Condition ($F_{2,46} = 34.720$, $p < 0.001$) was significant. Post hoc tests revealed significantly greater performance for the peak phase condition compared to the AM reference and the off-peak SIPI conditions, while the latter two conditions did not differ from each other.

F0 of 250 Hz. For the F0 of 250 Hz, the data were also first analyzed with a two-way RM-ANOVA using the factors SIPI phase and MD. The effect of SIPI phase ($F_{3,21} = 5.78$, $p = 0.005$) was significant, while neither MD ($F_{1,75} = 3.57$, $p = 0.073$) nor the interaction between these factors ($F_{3,21} = 0.42$, $p = 0.74$) was significant. Post hoc comparison showed significant differences between the peak phase (90°) and the off-peak phases (0° , 45° , 135°).

Then, to compare both peak- and off-peak phase conditions with the AM reference condition, we performed a two-way RM-ANOVA with the factors Condition (AM reference, peak phase, and off-peak phase, excluding 0° which was tested at MDs of 0.3 and 0.9 only) and MD (all MDs, except for the unmodulated condition). Note that in contrast to the F0 of 125 Hz, here, we would compare across all three conditions because the factor matrix was almost fully occupied. The main effect of Condition ($F_{2,70} = 21.79$,

$p < 0.001$) and the interaction with MD ($F_{8,70} = 3.2$, $p = 0.004$) were significant. The main effect of MD ($F_{4,70} = 2.37$, $p = 0.062$) was not significant. Post hoc mean comparisons showed that the peak phase condition was significantly better than the AM reference and off-peak SIPI phase conditions, and there was no difference between AM reference and off-peak SIPI phase conditions when pooling across MDs. However, in more detail, as suggested by the interaction between Condition and MD, at the MD of 0.1, both peak phase and off-peak phase conditions were significantly better than the AM reference condition, while at the other MDs, there were no significant differences between the three conditions (except for MD of 0.5). At MD of 0.5, the peak phase condition was significantly better than only the off-peak phase condition. In summary, as for the F0 of 125 Hz, the data at 250 Hz show significantly higher performance for the peak SIPI condition compared to the AM reference condition, but the relative improvement was larger at the lowest MD. The off-peak SIPI condition again showed no improvement at the higher MDs, but showed a comparable improvement to the peak phase SIPI condition at the lowest MD.

To compare the unmodulated pulse trains with the modulated pulse trains at the F0 of 250 Hz, a separate RM-ANOVA was performed using the factor Condition (AM reference, peak phase) and MD (all tested levels, including the unmodulated conditions). Both main factors, Condition ($F_{1,33} = 60.11$, $p < 0.001$) and MD ($F_{5,33} = 7.81$, $p < 0.001$) were significant, while their interaction was not significant ($F_{5,33} = 2.35$, $p = 0.062$). However, post hoc tests showed that the difference between AM reference and peak phase conditions was significant at MD of 0.1 only, and in the unmodulated conditions. Further, post hoc comparisons for the AM reference condition showed significantly higher performance for the MDs of 0.7, and 0.9 than for MD of 0.1.

Comparison Between F0s. Finally, effects observed at the two F0s were compared to each other. One caveat in this comparison is that the F0 of 250 Hz was tested in a separate test after the F0 of 125 Hz; thus, it is possible that learning effects favored the 250-Hz condition. Indeed, comparing the performance between the unmodulated SIPI conditions tested in the 125-Hz and 250-Hz blocks, thus, serving as control conditions, showed just significantly higher performance at 250 Hz (unpaired t test; $t(8) = 2.331$, $p = 0.048$). In the following analysis, the absolute effect of F0 should therefore be treated with caution while only interactions of the F0 with other factors, should be considered. Only the subset of four listeners who performed both F0 blocks were considered. A three-way RM-ANOVA was performed with the factors F0 (125 and 250 Hz), Condition (AM reference and peak

SIPI), and MD (all five levels) as well as their two- and three-way interactions. All main effects, F0 ($F_{1,97} = 5.9$, $p = 0.017$), Condition ($F_{1,97} = 95.41$, $p < 0.001$), and MD ($F_{4,97} = 11.77$, $p < 0.001$), were significant. Among the two-way interactions, only the interaction of MD and Condition ($F_{4,97} = 2.54$, $p = 0.045$) was significant (consistent with previous ANOVAs performed for the F0s separately). Importantly, none of the interactions involving F0 (neither two-way nor three-way) were significant. Further, a post hoc test for the AM reference condition only showed no significant difference between the two F0s. In summary, the absence of significant interactions involving the factor F0 suggests that the SIPI effect did not depend on the F0.

DISCUSSION

ITD Sensitivity in AM Reference Condition (Without SIPI Pulses)

The AM reference conditions showed increasing ITD sensitivity with increasing MD, with a saturation for MDs exceeding 0.5, particularly at the F0 of 250 Hz. This is consistent with studies showing decreasing ITD thresholds with increasing MD when testing NH listeners with high-frequency carriers and various types of AM (e.g., Bernstein and Trahiotis 2009; Klein-Hennig et al. 2011; Nuetzel and Hafter 1981). Note that our stimuli conveyed the ITD *both* in the carrier and the envelope (i.e., waveform ITD) whereas other studies presented ITD in the envelope alone. Still, such a mixed comparison seems to be justified because for carrier and modulation rates similar to ours, the ITD sensitivity is similar when the ITD is presented in the entire waveform versus the modulator alone (Noel and Eddington 2013).

Our effect of the MD is also in line with that found by Ihlefeld et al. (2014), who showed improving ITD sensitivity with increasing modulation depth tested in CI listeners with high-rate (1000 pps) pulse trains amplitude modulated at 100 Hz. Further, our results are consistent with those from Laback et al. (2011), who found increasing ITD sensitivity with increasing envelope pause time, a feature that co-varies with the MD.

Although we did not plan to study the effect of the F0 per se, we were surprised to observe better overall ITD sensitivity at 250 Hz than at 125 Hz, even if the effect did not reach significance for the AM reference condition alone. Bernstein and Trahiotis (2009) reported *lower* ITD sensitivity in NH listeners at a modulation frequency of 256 Hz compared to 128 Hz. In CI listeners, Noel and Eddington (2013) also showed lower ITD sensitivity for 1000-pps AM pulse trains with a modulation rate of 200 Hz compared to 100 Hz in all of the five listeners tested, although the

difference was surprisingly reported to be not significant. The evaluation of the overall effect of the F0s in the present study should, however, be treated with caution, because the two F0s were tested block-wise. Further, the control condition with SIPI pulses imposed on unmodulated pulse trains showed significantly better performance when tested within the second (250-Hz) block, thereby suggesting a training effect. Irrespective of this minor open question, the current data generally support previous findings that similar to NH listeners, CI listeners are sensitive to F0 modulation of both male and female speech.

General Effects of Introducing SIPI Pulses

In Srinivasan et al. (2018), it was shown that insertion of SIPI pulses into an unmodulated high-rate pulse train enhances ITD sensitivity. The current study extended the SIPI approach to modulated high-rate pulse trains, mimicking F0 modulation of voiced speech. While this modulation already provides a salient ITD cue (if the modulation depth is high), we found that adding SIPI pulses resulted in further enhancement of ITD sensitivity. Importantly, all conditions (with and without SIPI pulses) had been balanced in loudness, meaning that any increases in loudness caused by SIPI pulses had been accounted for by adjustments of overall pulse amplitude. Interestingly, SIPI-based improvements in ITD sensitivity were observed throughout the entire range of MDs tested. This means that the localization of speech stimuli that often have low MDs can strongly benefit from SIPI insertion.

The addition of a SIPI pulse can be thought of as changing the effective envelope, i.e., together with the preceding regular pulse, the SIPI pulse appears to be processed by the neural system similarly as a single pulse with an enhanced amplitude. Srinivasan et al. (2018) showed similar benefit in ITD sensitivity by introducing SIPI pulses with a low SIPI fraction and enhancing the amplitude of single pulses to adjust their short-time power to that of SIPI pulse pairs. This effect is, at first, independent of the presence of modulation in the pulse pattern, as shown by Srinivasan et al. (2018) and confirmed here for unmodulated pulse trains. For *modulated* pulse trains, we hypothesized that SIPI pulses further improve ITD sensitivity. Thus, we studied the contribution of SIPI pulses in addition to that of envelope ITD cues provided by external AM of pulse trains (here referred to as external envelope ITD cues). We indeed found a significant improvement of ITD sensitivity when adding SIPI pulses.

We further hypothesized that the SIPI effect would depend on the phase of the modulation cycle at which these SIPI pulses were inserted. This hypothesis

was based on previous evidence of greater contribution of amplitude information at the envelope peak in electric hearing to loudness (McKay and Henshall 2010) and to ITD (Hu et al. 2017). Indeed, we found a systematic dependence on the SIPI phase, with maximal SIPI-based effect at or around the envelope peak and minimal effect at off-peak phases. This result can be interpreted from three points of view.

First, because of the high amplitude at the envelope peak, SIPI pulses added at the envelope peak have the largest potential to enhance ITD sensitivity by means of their contribution to overall loudness (McKay and Henshall 2010) and by the overall level effect on ITD sensitivity in electric hearing (Egger et al. 2017). This interpretation is supported by the finding that at the F0 of 250 Hz, where both peak- and off-peak SIPI conditions were tested across all MDs, the off-peak condition was equally effective as the peak condition for the modulation depth of 0.1, where the amplitudes of peak and off-peak pulses differed only slightly. In contrast, at higher MDs where the off-peak amplitude was much lower than the peak amplitude, the off-peak SIPI pulses were not effective. This interpretation is also consistent with the lack of a SIPI-based benefit for pulse trains with SIPI pulse pairs having attenuated amplitudes (Srinivasan et al. 2018).

Second, the effect of the SIPI phase is potentially related to the readout window for extracting ITD information across the modulation cycle. For stimuli in the ITD-dominant frequency region in normal hearing, this window has been shown to be at the rising envelope segment (Dietz et al. 2013; Hu et al. 2017). In contrast, for high-frequency stimuli with a low-rate modulation in normal hearing (involving envelope ITD) and for low-rate pulse trains in electric hearing, the window appears to be located at/around the envelope peak (Hu et al. 2017). This peak dominance is consistent with our study: our listeners showed much better sensitivity when SIPI pulses were presented at/around the modulation peak compared to off-peak phases. Note that the comparison across studies is justified by the finding of Srinivasan et al. (2018) that periodically adding SIPI pulses to a high-rate pulse train with an ITD (as in the current study) has very similar effects as presenting corresponding low-rate pulse trains with the same ITD (as in Hu et al. 2017). Hu et al. (2017) attributed the peak dominance in electric hearing primarily to the integrative behavior of lateral superior olive (LSO) neurons which likely extract ITD in electric hearing. Note that one major difference is that they presented the ITD information at only a segment of the envelope, while we presented the ITD information in the whole envelope with added emphasis at the SIPI insertion phases. On a more detailed level, our results for the F0 of 125 Hz, where we tested phase steps with a finer resolution,

showed a significant asymmetry of the ITD sensitivity vs. SIPI phase pattern around the peak, with better sensitivity at the rising segment (67.5°) compared to the decaying segment (112.5°). While Hu et al. (2017) did not test their CI listeners at the decaying envelope segment, the asymmetric pattern in our study is consistent with their NH data using high-frequency stimuli and also with other NH studies on the weight of high-frequency envelope ITD across the modulation cycle (Hsieh et al. 2011; Klein-Hennig et al. 2011). Physiological studies on neural sensitivity to high-frequency envelope ITDs in barn owls (Nelson and Takahashi 2010) also report a stronger influence of the rising portion compared to the falling portion of the envelope on spike firing. A plausible explanation might be neural refractoriness reducing the response to a SIPI pulse pair following an envelope peak.

Third, the effect of the SIPI phase can be interpreted in terms of either sharpening or smearing the F0-based envelope modulation, thus either supporting or counteracting F0-based (external) envelope ITD cues. Inserting a SIPI pulse at the peak phase may be expected to enhance the envelope ITD cue because it increases the modulation depth and slope steepness (Klein-Hennig et al. 2011; Laback et al. 2011). In contrast, inserting a SIPI pulse at an off-peak phase may “fill” the modulation trough to some extent and therefore weaken the envelope ITD cue. Unfortunately, phase effects of SIPI pulses on the F0-based envelope ITD cue cannot easily be separated from their effects on loudness and the ITD readout window mentioned above, making it difficult to determine the relative contributions of these mechanisms from the current data. However, it is clear that the relative contribution of SIPI pulses to envelope sharpening/smearing will depend on the relative time constants of SIPI-evoked effective AM cues and of F0-modulation. Assuming a short-time constant of SIPI-evoked effective AM cues, the envelope sharpening/smearing effect can be expected to be stronger for higher F0s. Future studies, attempting to determine the contribution of different mechanisms, may test this prediction.

Finally, we hypothesized that SIPI pulses may be particularly effective for low modulation depths, where the F0-based (external) envelope ITD cues are weak. The data with the full set of six listeners tested at 125 Hz showed a constant SIPI-based benefit across MDs, thus not supporting the hypothesis. For the subset of four listeners tested at the F0 of 250 Hz, we actually found a larger benefit at the lowest MDs, but it is currently not clear if this effect can be generalized to the larger group tested at 125 Hz. Apart from this question, interestingly, our data generally show a SIPI-based improvement in ITD sensitivity throughout the range of MDs tested (up to 0.9). This suggests that SIPIs do not just restore the F0-

based envelope ITD cue, but instead they appear to provide an additional ITD cue.

Effect of F0 on SIPI-Based Benefit

Srinivasan et al. (2018) found that the benefit of adding SIPI pulses to unmodulated pulse trains declined when the SIPI rate exceeded 100 pps. To avoid such a rate limitation and to provide maximum benefit for ITD sensitivity, we fixed the SIPI rate to 62.5 Hz, which corresponded to one half and one quarter of the F0s of 125 and 250 Hz, respectively. Assuming that SIPI pulses evoke F0-independent highly stimulus-synchronized AN spikes and assuming reduced baseline ITD sensitivity at the F0 of 250 compared to 125 Hz, we conditionally hypothesized that our SIPI rate (fixed across F0s) would cause a SIPI-based benefit increasing with F0. In fact, results showed no reduction in baseline ITD sensitivity at the higher F0s (see the “ITD Sensitivity in AM Reference Condition (Without SIPI Pulses)” section) and, not surprisingly, no indication for higher SIPI-based benefit for the higher F0. Overall, we conclude that at least for a constant and low SIPI rate, SIPI insertion leads to similar improvements in ITD sensitivity at F0 frequencies ranging from male to female voices.

Comparison to Envelope Sharpening Approaches in the Literature

When inserted at the modulation peaks, SIPI pulses may sharpen F0-based envelope ITD cues (see the “General Effects of Introducing SIPI Pulses” section). Our approach can therefore be compared to other envelope sharpening approaches. Monaghan and Seeber (2016) proposed to enhance envelope ITD cues in reverberant signals by reducing the envelope level prior to direct-sound envelope peaks to the level of the preceding trough. Tested with NH listeners presented with an acoustic CI simulation, Monaghan and Seeber report an overall increase in ITD sensitivity, accompanied by a decrease in speech intelligibility for spatially close sources having a high direct-to-reverberant ratio. A modification limiting the zeroing to preserve the energy around the peaks avoided the deterioration for speech, with an unclear effect on the ITD efficiency. Francart et al. (2014) proposed an envelope enhancement strategy for bimodal hearing (acoustic in one ear and electric in the other ear). The strategy imposed a fast-rising modulation function on each F0 modulation peak of vowel-like stimuli and yielded a significant increase in ITD sensitivity.

While these approaches may be helpful in enhancing ITD cues of a single speaker in anechoic or echoic

environments, they may be harmful for configurations involving more than one speaker. The alteration of envelope information between F0 peaks of one speaker may distort envelope cues corresponding to the other speaker. Conversely, the strategic incorporation of SIPI pulses in a future CI stimulation strategy may reduce such distortions because each SIPI pulse modifies a small and temporally compact part of the envelope only.

SIPI pulses may also have an effect on F0-based temporal pitch perception in voiced speech. CI listeners are generally sensitive to temporal pitch cues, but the sensitivity declines rapidly with decreasing modulation depth (McKay et al. 1995; Vandali et al. 2013). This limits their access to pitch cues in real-life conditions in which the modulation depth is often reduced because of interfering sounds or reverberation. A recent study showed that SIPI pulses inserted at the envelope peak enhance pitch sensitivity particularly at low MDs (Lindenbeck et al. *in press*). Therefore, a future CI stimulation strategy implementing the SIPI approach has the potential to enhance both ITD and temporal pitch cues.

Note that increasing the amplitude of a pulse at the envelope peak may appear to be an attractive alternative to an extra (SIPI) pulse because it may be more energy efficient and it provides greater flexibility in pulse timing. However, compared to the SIPI approach, the amplitude increase required to evoke a comparable improvement in ITD sensitivity appears to yield a larger increase in loudness (Srinivasan et al. 2018). Thus, the preference for one or the other approach will probably depend on the particular coding requirements.

Generalization of Results

The bilateral CI listeners participating in this study were selected, based on the criteria of better-than-chance ITD sensitivity for unmodulated low-rate pulse trains and for unmodulated high-rate pulse trains with SIPI pulses. The first criterion served to increase the chances of obtaining interpretable data. The subset of CI listeners failing this criterion were not sensitive to even the most salient (i.e., low rate) ITD cues. Satisfying the second criterion of sensitivity to ITD cues in high-rate pulse trains with SIPI pulses ensured that listeners could be meaningfully tested on their comparative sensitivities to ITD cues in amplitude modulation without and with SIPI pulses. This raises the question to what extent the findings of the present study can be generalized to the whole population of bilateral CI listeners. The answer depends on the—currently unknown—proportion of CI listeners being ITD sensitive and the reason for the individual's lack of ITD sensitivity (see, e.g.,

Laback et al. 2015). For example, our results likely do not apply to listeners with lost ganglion cells in the relevant tonotopic region or with deprived binaural input during childhood (for the latter, see, e.g., Litovsky et al. 2010). However, our results are more likely generalizable to listeners showing a decline of ITD sensitivity because of deprivation of access to salient ITD cues as a consequence of using their clinical CI processors. It remains to be shown, though, if such an experience-based reduction of perceptual weights to ITD cues indeed occurs. Further, it needs to be shown if daily experience with salient ITD cues conveyed by a potential future clinical implementation of the SIPI approach or some variant, involving visual feedback for learning, indeed helps—even those individuals who would not pass our ITD-based inclusion criterion. Interestingly, one of our listeners (indicated with blue symbols in Fig. 3) showed close-to-chance performance in some conditions, but showed more sensitivity in other conditions (SIPI conditions at high modulation depths at the F0 of 125 Hz). This example shows that our ITD-based selection criterion allowed us to obtain useful results even in a listener with low overall ITD sensitivity.

SUMMARY AND CONCLUSIONS

We investigated the effects of inserting SIPI pulses into AM high-rate pulse trains on ITD sensitivity of bilateral CI listeners. The AM shapes and rates mimicked F0 modulation representing either male or female voiced speech. Stimuli were presented at a single, interaurally place-matched electrode pair. We hypothesized that introducing SIPI pulses at the peak of the AM cycle improves ITD sensitivity, especially at low MDs. We therefore systematically varied the MD (across a wide range) and the SIPI phase (across the positive cycle of the AM period) and tested ITD sensitivity in a left/right discrimination task for both F0s. We found that SIPI pulses inserted at or around F0-based envelope peaks enhanced ITD sensitivity relative to the AM conditions without SIPI insertion. On a more detailed level, SIPI insertion at the late rising segment of the modulation cycle was more effective than at the early decaying segment, which was shown to be consistent with normal hearing data from the literature using high-frequency stimuli. SIPI pulses inserted further off the modulation peak (in both directions), however, yielded no advantage in ITD sensitivity. Interestingly, SIPI-based benefit was observed throughout the range of modulation depths tested and for both F0s representing either male or female voiced speech. Together with the recent results of the beneficial

effects of SIPI pulses on temporal pitch sensitivity (Lindenbeck et al. *in press*), the current data suggest that the SIPI approach has the potential to encode salient ITD and pitch cues at relatively high pulse rates that are required to maintain high speech understanding. Better access to ITD and pitch cues are then expected particularly under real-life conditions involving background noise or reverberation, both of which reduce the effective modulation depth.

ACKNOWLEDGMENTS

We thank all our listeners for their participation in our study. We thank Michael Mihocic for assistance with data collection and the experimental hardware and software setup. We thank the Institute of Ion Physics and Applied Physics, Leopold-Franzens-University of Innsbruck, Austria, for providing the equipment for direct electric stimulation. We thank the associate editor and two anonymous reviewers for helpful comments on an earlier version of this article.

Funding Information This study was funded by the National Institutes of Health grant R01 DC 005775 via a subcontract. Additional funding was provided by the European Commission (Project ALT, Grant 691229) and the Danube Partnership project (MULT-DR 11/2017).

COMPLIANCE WITH ETHICAL STANDARDS

All procedures presented here were approved by the ethics committee of the Medical University of Vienna (vote no. 2155/2013).

Conflict of Interest The authors declare that they have no conflict of interest.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES

- ARORA K, DAWSON P, DOWELL R, VANDALI A (2009) Electrical stimulation rate effects on speech perception in cochlear implants. *Int J Audiol* 48:561–567
- BERNSTEIN LR, TRAHOTIS C (2002) Enhancing sensitivity to interaural delays at high frequencies by using “transposed stimuli”. *J Acoust Soc Am* 112:1026–1036
- BERNSTEIN LR, TRAHOTIS C (2009) How sensitivity to ongoing interaural temporal disparities is affected by manipulations of temporal features of the envelopes of high-frequency stimuli. *J Acoust Soc Am* 125:3234–3242
- BRUGHERA A, DUNAI L, HARTMANN WM (2013) Human interaural time difference thresholds for sine tones: the high-frequency limit. *J Acoust Soc Am* 133:2839–2855
- BUECHEL BD, HANCOCK KE, CHUNG Y, 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
- CHURCHILL TH, KAN A, GOUELL MJ, LITOVSKY RY (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
- DIETZ M, MARQUARDT T, SALMINEN NH, McALPINE D (2013) Emphasis of spatial cues in the temporal fine structure during the rising segments of amplitude-modulated sounds. *Proc Natl Acad Sci U S A* 110:15151–15156
- DORMAN MF, WILSON BS (2004) The design and function of cochlear implants: fusing medicine, neural science and engineering, these devices transform human speech into an electrical code that deafened ears can understand. *Am Sci* 92(5):436–445
- EGGER K, MAJDAK P, LABACK B (2016) Channel interaction and current level affect across-electrode integration of interaural time differences in bilateral cochlear-implant listeners. *J Assoc Res Otolaryngol* 17:55–67
- EGGER K, MAJDAK P, 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, LENSSEN A, WOUTERS J (2014) Modulation enhancement in the electrical signal improves perception of interaural time differences with bimodal stimulation. *J Assoc Res Otolaryngol* 15:633–647
- GAROFOLO JS, LAMEL LF, FISHER WM, FISCUS JG, PALLETT DS (1993) DARPA TIMIT acoustic-phonetic continuous speech corpus CD-ROM. NIST speech disc 1-1.1. NASA STIRcon Tech Rep N. 93
- GRANTHAM DW, ASHMEAD DH, RICKETTS TA, HAYNES DS, LABADIE RF (2008) Interaural time and level difference thresholds for acoustically presented signals in post-lingually deafened adults fitted with bilateral cochlear implants using CIS+ processing. *Ear Hear* 29:33–44
- GRANTHAM DW, ASHMEAD DH, RICKETTS TA, LABADIE RF, HAYNES DS (2007) Horizontal-plane localization of noise and speech signals by postlingually deafened adults fitted with bilateral cochlear implants. *Ear Hear* 28:524–541
- GREEN DM, SWETS JA (1966) Signal detection theory and psychophysics. Wiley New York
- HANCOCK KE, CHUNG Y, DELGUTTE B (2012) Neural ITD coding with bilateral cochlear implants: effect of binaurally coherent jitter. *J Neurophysiol* 108:714–728
- HOCHMAIR I, NOPP P, JOLLY C, SCHMIDT M, SCHÖSSER H, GARNHAM C, ANDERSON I (2006) MED-EL cochlear implants: state of the art and a glimpse into the future. *Trends Amplif* 10:201–219
- VAN HOESEL R, BÖHM M, PESCH J, VANDALI A, BATTMER RD, LENARZ T (2008) Binaural speech unmasking and localization in noise with bilateral cochlear implants using envelope and fine-timing based strategies. *J Acoust Soc Am* 123:2249–2263
- VAN HOESEL RJM, JONES GL, LITOVSKY RY (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
- Hsieh I-H, PETROSYAN A, GONÇALVES ÓF, HICKOK G, SABERI K (2011) Observer weighting of interaural cues in positive and negative envelope slopes of amplitude-modulated waveforms. *Hear Res* 277:143–151
- HU H, EWERT SD, McALPINE D, 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
- IHLEFELD A, KAN A, LITOVSKY RY (2014) Across-frequency combination of interaural time difference in bilateral cochlear implant listeners. *Front Syst Neurosci* 8:1–10

- IHLEFELD A, LITOVSKY RY (2012) Interaural level differences do not suffice for restoring spatial release from masking in simulated cochlear implant listening. *PLoS One* 7:e45296
- JESTEADT W (1980) An adaptive procedure for subjective judgments. *Percept Psychophys* 28:85–88
- KAN A, LITOVSKY RY, GOUELL MJ (2015) Effects of interaural pitch matching and auditory image centering on binaural sensitivity in cochlear implant users. *Ear Hear* 36(3):e62
- KIDD G JR, MASON CR, BEST V, MARRONE N (2010) Stimulus factors influencing spatial release from speech-on-speech masking. *J Acoust Soc Am* 128:1965–1978
- KLEIN SA (2001) Measuring, estimating, and understanding the psychometric function: a commentary. *Percept Psychophys* 63:1421–1455
- KLEIN-HENNIG M, DIETZ M, HOHMANN V, EWERT SD (2011) The influence of different segments of the ongoing envelope on sensitivity to interaural time delays. *J Acoust Soc Am* 129:3856–3872
- KLUMPP RG, EADY HR (1956) Some measurements of interaural time difference thresholds. *J Acoust Soc Am* 28:859–860
- KOHLER KJ (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*. pp. 1938–1941. IEEE
- LABACK B, EGGER K, MAJDAK P (2015) Perception and coding of interaural time differences with bilateral cochlear implants. *Hear Res* 322:138–150
- LABACK B, 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, 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 WA, 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, 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
- LEVITT H (1971) Transformed up-down methods in psychoacoustics. *J Acoust Soc Am* 49:467–477
- LINDENBECK M, LABACK B, MAJDAK P, SRINIVASAN S (IN PRESS) TEMPORAL-PITCH SENSITIVITY IN ELECTRIC HEARING WITH AMPLITUDE MODULATION AND SHORT INTER-PULSE INTERVALS. *J ACOUST SOC AM*
- LITOVSKY RY, JONES GL, AGRAWAL S, VAN HOESEL R (2010) Effect of age at onset of deafness on binaural sensitivity in electric hearing in humans. *J Acoust Soc Am* 127(1):400–414
- LOIZOU PC, POROY O, DORMAN M (2000) The effect of parametric variations of cochlear implant processors on speech understanding. *J Acoust Soc Am* 108:790–802
- MACPHERSON EA, MIDDLEBROOKS JC (2002) Listener weighting of cues for lateral angle: the duplex theory of sound localization revisited. *J Acoust Soc Am* 111:2219–2236
- MAJDAK P, LABACK B, BAUMGARTNER W-D (2006) Effects of interaural time differences in fine structure and envelope on lateral discrimination in electric hearing. *J Acoust Soc Am* 120:2190–2201
- McKAY CM, HENSHALL KR (2010) Amplitude modulation and loudness in cochlear implantees. *J Assoc Res Otolaryngol* 11:101–111
- McKAY CM, McDERMOTT HJ (1998) Loudness perception with pulsatile electrical stimulation: the effect of interpulse intervals. *J Acoust Soc Am* 104:1061–1074
- McKAY CM, McDERMOTT HJ, CLARK GM (1995) Pitch matching of amplitude-modulated current pulse trains by cochlear implantees: the effect of modulation depth. *J Acoust Soc Am* 97:1777–1785
- MIDDLEBROOKS JC, ONSAN ZA (2012) Stream segregation with high spatial acuity. *J Acoust Soc Am* 132:3896–3911
- MONAGHAN JJ, SEEGER BU (2016) A method to enhance the use of interaural time differences for cochlear implants in reverberant environments. *J Acoust Soc Am* 140:1116–1129
- NELSON BS, TAKAHASHI TT (2010) Spatial hearing in echoic environments: the role of the envelope in owls. *Neuron*. 67:643–655
- NOEL VA, EDDINGTON DK (2013) Sensitivity of bilateral cochlear implant users to fine-structure and envelope interaural time differences. *J Acoust Soc Am* 133:2314–2328
- NUETZEL JM, HAFTER ER (1981) Discrimination of interaural delays in complex waveforms: spectral effects. *J Acoust Soc Am* 69:1112–1118
- 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. *Speech Prosody* 7:305–309
- POON BB, EDDINGTON DK, NOEL V, COLBURN HS (2009) Sensitivity to interaural time difference with bilateral cochlear implants: development over time and effect of interaural electrode spacing. *J Acoust Soc Am* 126:806–815
- RISS D, HAMZAVI J-S, BLINEDER M, HONEDER C, EHRENREICH I, KAIDER A, BAUMGARTNER W-D, GSTOETTNER W, ARNOLDNER C (2014) FS4, FS4-p, and FSP: a 4-month crossover study of three fine structure sound-coding strategies. *Ear Hear* 35:e272–e281
- SEEGER BU, FASTL H (2008) Localization cues with bilateral cochlear implants. *J Acoust Soc Am* 123:1030–1042
- SHANNON RV (1985) Threshold and loudness functions for pulsatile stimulation of cochlear implants. *Hear Res* 18:135–143
- SMITH Z (2010) Improved sensitivity to interaural time differences with the FAST coding strategy. 11th International Conference on Cochlear Implants and Other Implantable Auditory Technologies. Stockholm, Sweden
- SRINIVASAN S, LABACK B, MAJDAK P, DELGUTTE B (2018) Introducing short interpulse intervals in high-rate pulse trains enhances binaural timing sensitivity in electric hearing. *J Assoc Res Otolaryngol* 19:301–315
- STRUTT JW (1876) Our perception of the direction of a source of sound. *Nature*. 14:32–33
- THAKKAR T, KAN A, JONES HG, LITOVSKY RY (2018) Mixed stimulation rates to improve sensitivity of interaural timing differences in bilateral cochlear implant listeners. *J Acoust Soc Am* 143:1428–1440
- VANDALI A, SLY D, COWAN R, VAN HOESEL R (2013) Pitch and loudness matching of unmodulated and modulated stimuli in cochlear implantees. *Hear Res* 302:32–49
- WIGHTMAN FL, KISTLER DJ (1992) The dominant role of low-frequency interaural time differences in sound localization. *J Acoust Soc Am* 91:1648–1661
- WILLIGES B, JÜRGENS T, HU H, DIETZ M (2018) Coherent coding of enhanced interaural cues improves sound localization in noise with bilateral cochlear implants. *Trends Hear* 22:2331216518781746
- WILSON BS, FINLEY CC, LAWSON DT, WOLFORD RD, EDDINGTON DK, RABINOWITZ WM (1991) Better speech recognition with cochlear implants. *Nature*. 352:236–238
- WILSON BS (2006) Speech processing strategies. In: Cooper HR, Craddock LC (eds) *Cochlear implants: a practical guide*, 2nd edn. John Wiley & Sons, Hoboken, NJ, pp 21–69