Spectrotemporal weighting of binaural cues: Effects of a diotic interferer on discrimination of dynamic interaural differences

Jacqueline M. Bibee

*Department of Speech and Hearing Sciences, University of Washington, 1417 Northeast 42nd Street,
Seattle Washington 98105, USA*

G. Christopher Stecker^{a)}

*Department of Hearing and Speech Sciences, Vanderbilt University School of Medicine,
1215 21st Avenue South, Nashville Tennessee 37232, USA*

(Received 27 April 2016; revised 23 September 2016; accepted 29 September 2016; published online 13 October 2016)

Spatial judgments are often dominated by low-frequency binaural cues and onset cues when binaural cues vary across the spectrum and duration, respectively, of a brief sound. This study combined these dimensions to assess the spectrotemporal weighting of binaural information. Listeners discriminated target interaural time difference (ITD) and interaural level difference (ILD) carried by the onset, offset, or full duration of a 4-kHz Gabor click train with a 2-ms period in the presence or absence of a diotic 500-Hz interferer tone. ITD and ILD thresholds were significantly elevated by the interferer in all conditions and by a similar amount to previous reports for static cues. Binaural interference was dramatically greater for ITD targets lacking onset cues compared to onset and full-duration conditions. Binaural interference for ILD targets was similar across dynamic-cue conditions. These effects mirror the baseline discriminability of dynamic ITD and ILD cues [Stecker and Brown. (2010). *J. Acoust. Soc. Am.* **127**, 3092–3103], consistent with stronger interference for less-robust/higher-variance cues. The results support the view that binaural cue integration occurs simultaneously across multiple variance-weighted dimensions, including time and frequency.

© 2016 Acoustical Society of America. [<http://dx.doi.org/10.1121/1.4964708>]

[LRB]

Pages: 2584–2592

I. INTRODUCTION

Spatial hearing by human and animal listeners involves auditory spatial cues that vary across several dimensions, including cue type [e.g., interaural time difference (ITD) and interaural level difference (ILD)], frequency, and time. These cues arise acoustically and are therefore subject to acoustical limitations and distortions. For example, ILD cues are generally more robust at high than low frequencies, ITD cues can be ambiguous at high frequencies, and both cues may be significantly distorted by echoes and reverberation (Stecker and Gallun, 2012). Accurate sound localization therefore requires that these various cues—which may all be present but not necessarily in “agreement” in a given stimulus—be integrated in a manner that is sensitive to their relative reliability. That is, the perceptual weighting of auditory spatial cues should favor robust and reliable cues over weaker or unreliable cues.

Clear examples of perceptual weighting and binaural-cue dominance have been demonstrated across each of these dimensions (cue type, frequency, and time). Weighting across cue type is demonstrated in ITD/ILD “trading ratios” that place greater weight on ITD at low vs high narrowband frequencies (David *et al.*, 1959; Harris, 1960) and slow vs rapid modulation rates (Stecker, 2010). Weighting across frequency is demonstrated in studies of “binaural interference” that

reveal the dominance of low- over high-frequency ITD cues and high- over low-frequency ILD cues (McFadden and Pasanen, 1976; Heller and Trahiotis, 1995; Heller and Richards, 2010). Weighting over time is demonstrated by the dominance of onset over later-arriving binaural cues in low-frequency pure tones (Stecker and Bibee, 2014; Diedesch and Stecker, 2015) and rapidly modulated sounds (Haftner and Dye, 1983; Saberi, 1996; Freyman *et al.*, 1997; Stecker, 2014). Each of these examples appears consistent with the dominance of more reliable over less reliable cues.

The current study aimed to investigate binaural-cue dominance across two dimensions simultaneously (frequency and time) by measuring binaural interference under conditions of onset dominance. Specifically, we set out to measure the effects of a low-frequency diotic interferer on the discrimination of high-frequency target ITD and ILD cues available primarily in the highly weighted onset vs the weakly weighted offset portion of a brief sound. If binaural interference reflects the dominance of normally reliable cues over normally less reliable cues, we hypothesize that greater interference should be observed for target cues presented later in the sound than at sound onset. That is, we asked whether the robust onset cue would be less susceptible to binaural interference than would the later, ongoing cues.

Our approach borrowed strongly from that of Stecker and Brown (2010), who measured ITD and ILD thresholds in 4000-Hz Gabor click trains (narrowband-filtered impulse trains) in three target conditions. A static-cue condition (condition “RR”), in which each of 16 clicks carried the same

^{a)}Electronic mail: cstecker@spatialhearing.org

rightward ITD or ILD, was contrasted with 2 dynamic-cue conditions. Dynamic cues either started diotically and grew to a large rightward value at sound offset (condition “OR”), or followed the opposite temporal pattern (condition “R0”). [Stecker and Brown \(2010\)](#) measured binaural thresholds adaptively across a range of interclick intervals (ICIs) from 2 to 10 ms (click rates from 100 to 500 Hz). At the shortest tested ICI (2 ms), [Stecker and Brown \(2010\)](#) found similar ITD thresholds in conditions RR and R0 (both of which featured interaural onset cues), but significantly elevated thresholds in condition OR, which lacked onset cues. The result was consistent with strong onset dominance at high rates in that listeners were clearly impaired when the onset cue was not available. For ILD targets, [Stecker and Brown \(2010\)](#) found no difference between conditions R0 and OR. In the current study, we attempted to replicate the 2-ms ICI threshold measurements of [Stecker and Brown \(2010\)](#) while adding an additional variable: The presence or absence of a simultaneous 500-Hz diotic interfering tone. The hypothesis was that ITD thresholds in conditions RR and R0 would be relatively unaffected by the interferer because of the robust onset cue, whereas thresholds in condition OR (which omits the onset cue) would be significantly elevated by the interferer.

Several studies have described weaker binaural interference for temporal ([Trahiotis and Bernstein, 1990](#); [Woods and Colburn, 1992](#); [Hill and Darwin, 1996](#)) and harmonic ([Buell and Hafer, 1991](#); [Hill and Darwin, 1996](#)) relationships that reduce the perceptual fusion of targets and interferers, or with sequential grouping cues ([Best et al., 2007](#)) that do the same. Here, we sought to maintain fusion of the target and interferer in two ways. First, the target and interferer were gated on and off simultaneously. Second, the interferer was always presented at the fundamental frequency of the target complex (500 Hz $\pm$ 10% synchronized roving).

II. METHODS

The stimuli and design of the study were modeled closely on those of [Stecker and Brown \(2010\)](#). Behavioral testing took place in the University of Washington Speech and Hearing Clinic, Seattle, WA. All procedures, including recruiting, consenting, and testing of human subjects, followed guidelines of the University of Washington Human Subjects Division and were reviewed and approved by the cognizant Institutional Review Board.

A. Participants

Eight adults aged 24–40 yr (two males) participated in the experiments. One participant (0510) was the first author, and another (0513) was a rotation student working on other projects in the laboratory. Others were paid participants naive to the purpose of the study. All participants demonstrated normal hearing levels [<20 dB hearing level (HL) from 250 Hz to 8 kHz and <10 dB threshold difference between ears], as documented by an audiogram on the first day of testing or within the previous year, and negative history of neurological disorder. One participant (1118) who

completed the study was omitted from analyses due to ILD thresholds that exceeded the group mean by $>3\sigma$ across all conditions and did not improve with training.

B. Stimuli

As in the study of [Stecker and Brown \(2010\)](#), target stimuli were trains of 16 Gabor clicks (Gaussian-windowed tone bursts). Each click consisted of a 4 kHz cosine carrier multiplied by a Gaussian temporal envelope with $\sigma = 221 \mu\text{s}$. The result is a narrowband-filtered impulse with center frequency 4 kHz and half-maximal bandwidth of 1.8 kHz. Trains of 16 such clicks were synthesized at an ICI of 2 ms and presented at a sampling rate of 48.828 kHz (Tucker-Davis Technologies RP2.1, Alachua, FL) via circumaural electrostatic earphones (Stax SR-307, Saitama, Japan). The average peak-equivalent binaural level was 80 dB sound pressure level (SPL).

In separately tested conditions, target stimuli carried either an ITD—the value of which is denoted Δt —that led in the right ear, or an ILD—denoted ΔL —that presented greater intensity to the right ear. Binaural cues were presented in three temporal configurations as previously defined by [Stecker and Brown \(2010\)](#): In condition RR, ITD or ILD was constant at Δt or ΔL for all clicks. In condition R0, the ITD or ILD was equal to Δt or ΔL on the first click and then decreased linearly to zero on the final click. In condition OR, the reverse was true; ITD or ILD increased linearly from zero on the first click to Δt or ΔL on the last. Figure 1 illustrates ITD target stimuli in all three conditions. Reference stimuli were identical to targets but carried 0 μs ITD and 0 dB ILD on all clicks. Target ICI varied randomly (“roved”) by $\pm 10\%$ across intervals to minimize listeners’ access to pitch cues in the dynamic-ITD conditions (OR and R0).

The interferer, when present, was a diotic (zero ITD and ILD) sinusoidal tone at 80 dB peak-equivalent SPL. Interferer frequency was nominally 500 Hz but roved $\pm 10\%$ across intervals in tandem with 1/ICI of the reference and target stimuli. Thus, the interferer tone always matched the fundamental frequency of the Gabor click train. Total duration of the interferer was 32 ms (matching the target and reference stimuli), including raised-cosine rise/fall ramps of 10 ms each. The interferer waveform is illustrated in the bottom panel of Fig. 1. In interference conditions, interferers were presented synchronously with both reference and target stimuli. In no-interference conditions, reference and target stimuli were presented on their own.

C. Procedures

Binaural discrimination was assessed using a four-interval, two-alternative, forced choice (4I2AFC) task. The right-leading target was presented randomly in the second or third interval; the other three intervals presented diotic reference stimuli. Interference conditions presented the interferer simultaneously with reference and target stimuli in all four intervals. Participants indicated by button-press whether the second or third interval contained the right-leading target stimulus. Feedback indicated the target interval by illuminating the

FIG. 1. Stimuli employed in the study. Target and reference stimuli were trains of 16 Gabor clicks (digitally filtered impulses spanning 3–5 kHz) repeating at an ICI of 2 ms. Target intervals carried ITD or ILD, depending on condition; reference intervals were diotic. In the ITD condition (shown), target ITD was either constant at Δt over the train duration (condition RR, top), diminished from Δt at sound onset to $0 \mu\text{s}$ at sound offset (condition R0, second from top), or increased from $0 \mu\text{s}$ at onset to Δt at offset (condition OR, third from top). Target ILD (not shown) was applied in a temporally similar fashion for the ILD conditions. Binaural interference conditions additionally presented a diotic 500-Hz tonal interferer (“int,” bottom). The interferer had a total duration of 32 ms (equal to the target) including 10 ms raised-cosine rise/fall ramps and was presented synchronously with the target. Both the targets and interferers were presented at 80 dB peak-equivalent SPL.

corresponding light emitting diode (LED) immediately following each trial.

Thresholds were obtained using a two-down one-up adaptive procedure tracking 71% correct. Two simultaneous tracks were interleaved within each run. One of the two tracks was chosen randomly on each trial. The tracked parameter was Δt or ΔL . In ITD conditions, Δt started each run at $500 \mu\text{s}$ and was adjusted in base-10 logarithmic steps of 0.2 (i.e., 58% change) until four adaptive reversals were recorded and then in steps of 0.05 (12%) for an additional eight reversals. Values were limited to the range $\pm 600 \mu\text{s}$. The ITD threshold for each track was computed as the geometric mean of the final eight reversals. In ILD conditions, ΔL started each run at 10 dB and was adjusted in steps of 2 dB for two reversals, 0.5 dB for four additional reversals, and 0.1 dB for the final eight reversals. Values were limited to the range ± 12 dB. ILD thresholds for each track were computed as the arithmetic mean of the final eight reversals.

Testing began with several practice runs presenting both ITD and ILD targets in the RR configuration. Listeners completed at least four such practice runs with no interferer, and at least two runs with interferer present, before continuing. Following practice runs, each participant completed at least 8 tracks (4 runs of approximately 100 trials each) for each combination of cue type (ITD or ILD), configuration (RR, R0, or OR), and interference condition (interferer present or absent). Testing order was randomized in a counterbalanced fashion. A fifth run was completed in any case where one or two adaptive tracks failed to converge asymptotically. Most often, this occurred when thresholds appeared near the maximum of the tracking range (12 dB or $600 \mu\text{s}$). The adaptive threshold for each condition was defined as the geometric mean threshold ITD or the arithmetic mean threshold ILD of at least seven successful tracks per participant. No adaptive

threshold was recorded if fewer than seven tracks converged asymptotically.

Ceiling effects in some experimental conditions led to a large number of “no threshold” conditions, prompting a second analysis of the data. Average thresholds were computed by fitting a Weibull psychometric function to the full set of trials obtained for that combination of listener and condition and extrapolating the 71% correct point using MATLAB functions FitWeibTAFc and FindThreshWeibTAFc from the Psychophysics Toolbox extensions (version 3.0.10; Brainard, 1997). Psychometric fits provided more stable threshold estimates in conditions where the adaptive ceiling artificially limited the collection of adaptive reversals. Fit thresholds were computed independently of adaptive thresholds for all subjects and conditions. Statistical comparisons of fit thresholds across conditions were made in the same way as for adaptive thresholds.

D. Statistical data analysis

Adaptively determined thresholds for each participant and condition are plotted as symbols in Fig. 2. Also plotted are sampling distributions of the across-subject mean threshold in each condition, estimated by 5000-fold bootstrapped sampling of the geometric mean ITD threshold or arithmetic mean ILD threshold using MATLAB’s bootstrap function. Bootstrap tests were also used for paired comparisons across conditions (one-tailed tests of the $0\text{R} > \text{R0} > \text{RR}$ difference reported previously by Stecker and Brown, 2010). For paired tests, we resampled (5000 times) the across-subject mean of differences between conditions and computed the proportion of bootstrapped estimates exhibiting ≤ 0 difference, i.e., the raw p -value, expressed to one significant digit (Efron and Tibshirani, 1986; Fox, 2008; Stecker and Bibee, 2014).

A binaural interference index (BII) was computed for each combination of subject, cue type, and cue configuration (Fig. 3). For ITD, this was defined as the base-two log ratio of ITD thresholds obtained with interferer present vs absent. For ILD, BII was defined as the threshold difference (in dB) between interferer and no-interferer conditions. In both cases, larger positive numbers indicate greater interference, and $\text{BII} = 0$ indicates a lack of interference.

Statistical comparisons of BII values utilized 5000-fold bootstrap tests addressing the null hypothesis of no interference ($\text{BII} \leq 0$) in each condition. Similar paired bootstrap tests assessed BII differences across conditions (e.g., $0\text{R} - \text{RR} \leq 0$). For both types of tests, MATLAB’s bootstrap function was used to estimate the sampling distribution of the mean BII or BII difference across 5000 iterations. Results of these tests are indicated in the text as raw p -values expressing the proportion of bootstrapped estimates consistent with the null hypothesis to one significant digit (e.g., “ $p < 0.002$ ”).

III. RESULTS AND DISCUSSION

A. Threshold differences in no-interference conditions

The effects of binaural cue configuration (RR vs R0 vs OR) when the interferer was absent are summarized along the horizontal axes of Fig. 2. In each panel, symbols plot

FIG. 2. Threshold ITD (left) and ILD (right), determined adaptively and plotted for conditions with interferer present (vertical axis) vs conditions with interferer absent (horizontal axis). Individual subjects' mean thresholds (across runs) are plotted by symbol type, with conditions RR, R0, and OR appearing in black, gray, and white, respectively. Symbols falling above the black unity line indicate threshold elevation in the presence of a diotic 500-Hz interferer. Note ceiling effects (NT = "no threshold") which were especially prevalent for ITD thresholds measured for OR targets with interferer present. Curves along each axis plot the bootstrap-estimated sampling distributions of the cross-subject mean in each condition; black: RR, gray: R0, dashed: OR. For reference, group-mean data measured by [Stecker and Brown \(2010\)](#) for identical target stimuli presented without interferers are plotted as stars along the horizontal axis.

ITD and ILD thresholds of individual listeners. Curves along each axis plot bootstrap-estimated sampling distributions of the across-subject mean in each condition. When no interferer was present, ITD thresholds were lower in RR and R0 (which themselves did not significantly differ, $p < 0.2$) than OR (vs RR: $p < 0.0002$; vs R0: $p < 0.0002$). The result is

consistent with the pattern reported by [Stecker and Brown \(2010\)](#); means from that study are plotted as stars in Fig. 2) and with reduced sensitivity to post-onset ITD at short ICI. Also consistent with [Stecker and Brown \(2010\)](#), ILD thresholds were best in RR (vs R0: $p < 0.0002$; vs OR: $p < 0.0002$) but did not differ significantly between R0 and

FIG. 3. BII plotted against cue configuration condition for individual subjects (symbols and legend as in Fig. 2). The mean across subjects is indicated by a large black circle and horizontal line in each panel. For ITD (left), BII is the log ratio of threshold ITD determined with interferer present vs absent. For ILD (right), BII is defined as the dB difference in threshold determined with interferer present vs absent. (a) BII computed from adaptive-tracked thresholds as in Fig. 2. Upward-pointing arrows indicated values that could not be reliably estimated due to unmeasurable thresholds (NT in Fig. 2); in all but two cases, thresholds were obtainable in the no-interferer condition but not when the interferer was present. (b) BII computed from psychometric-function fits to trial data. Asterisks (*) indicate significant differences in mean BII across conditions ($p < 0.03$ by 5000-fold paired bootstrap test).

0R ($p < 0.06$). These results suggest that ILD cues are processed at both onset and offset (Stecker and Brown, 2012).

B. Threshold differences in interference conditions

The vertical axis of Fig. 2 plots ITD and ILD thresholds obtained in the presence of the diotic interferer. The effects of binaural interference are thus illustrated by comparing threshold values to values on the horizontal axis (no interferer). For both cue types and all three cue configurations, symbols appear predominantly above the diagonal, consistent with threshold elevations in the presence of the diotic low-frequency interferer. The degree of interference was further quantified by the BII and plotted in Fig. 3. Large symbols plot mean values across subjects. Figure 3(a) plots BII values computed from adaptively determined thresholds. Upward arrows in Fig. 3 indicate values that could not be appropriately estimated due to ceiling effects on adaptively determined thresholds [NT in Fig. 2(a)]; the plotted values in these cases substitute the value $500\ \mu\text{s}$ for the undetermined thresholds. The proportion of subjects affected in this way was significant for ITD condition 0R (seven out of seven, $p < 0.05$ via sign test) but not in other conditions. Figure 3(b) plots BII values estimated via psychometric fits to trial data. Psychometric fits more clearly quantified the degree of binaural interference in these cases, with thresholds falling well beyond the ceiling of the adaptive procedure and typically exceeding the no-interferer threshold by a factor of eight or more ($\text{BII} > 3$). Bootstrap tests indicated statistically significant interference ($\text{BII} > 0$) in all conditions for both ITD (RR: $p < 0.0002$, R0: $p < 0.02$, 0R: $p < 0.002$) and ILD (RR: $p < 0.0002$, R0: $p < 0.0002$, 0R: $p < 0.02$).

C. Greater interference for ITD targets lacking onset cues

Due to the large number of cases in which ITD thresholds could not be measured adaptively with the interferer present, we could not make meaningful quantitative comparisons of BII calculated from adaptive ITD thresholds [Fig. 3(a)]. However, comparisons of psychometric-fit BII values [Fig. 3(b)] for ITD targets were possible. These revealed significantly greater interference in condition 0R than R0 ($p < 0.02$, paired bootstrap test) but no difference between R0 and RR ($p < 0.5$). A majority of participants exhibited dramatic threshold elevations in condition 0R (which lacks onset ITD cues) relative to other conditions. As a result, mean BII in condition 0R was equal to 1.90 (i.e., a quadrupling of ITD threshold). Mean BII in other conditions (1.07 in RR and 0.72 in R0) indicated an approximate doubling of thresholds by interference, roughly on par with previous reports of binaural interference with high-frequency ITD targets and low-frequency diotic interferers (McFadden and Pasanen, 1976; Heller and Richards, 2010).

D. Similar interference for target ILD presented at sound onset or offset

Comparing BII computed from adaptively determined ILD thresholds [Fig. 3(a)] across conditions revealed on

average 0.7, 1.3, and 0.9 dB of interference in conditions RR, R0, and 0R, respectively. BII computed from psychometric-fit ILD thresholds [Fig. 3(b)] revealed slightly weaker interference (0.5, 1.1, and 0.6 dB in RR, R0, and 0R, respectively). In comparison, Heller and Richards (2010) reported roughly 0.2 dB threshold elevation for high-frequency ILD targets and low-frequency diotic interferers, using narrow bands of noise. Regardless of the type of threshold calculation, significantly greater interference was observed in condition R0 than RR ($p < 0.01$) or 0R ($p < 0.03$), a result that runs counter to the hypothesis of reduced interference for cues at sound onset. Instead, the results suggest that late-arriving cues continue to play an important role in ILD discrimination (Stecker and Brown, 2012), even in the presence of an effective diotic interferer. BII did not differ significantly between condition 0R and RR ($p < 0.4$).

E. Individual differences in binaural discrimination and binaural interference

Visual inspection of Figs. 2 and 3 suggests a range of psychophysical performance in the current study, consistent with several previous studies (Woods and Colburn, 1992; Heller and Trahiotis, 1995; Best *et al.*, 2007). Some listeners appeared minimally affected by the interferer in many conditions, while others demonstrated large thresholds and/or inability to perform the task in some interferer conditions. Substantial variation in ITD thresholds was observed across subjects, spanning at least 100–300 μs in the RR-no-interferer condition, and a greater range in other conditions. That degree of variability is roughly in line with previous reports of envelope-ITD discrimination at short ICI (Stecker and Brown, 2010). However, ITD thresholds were significantly higher in the current study as compared to those reported by Stecker and Brown (2010) for the same task (RR: $p < 0.0002$, R0: $p < 0.02$, 0R: $p < 0.0002$). Although the difference might simply reflect variation in listeners' intrinsic sensitivity, procedural differences could potentially play an additional role. Stecker and Brown (2010) tested ITD and ILD conditions in separate experiments (i.e., cue type was consistent from run to run) and did not present an interferer. Stimuli in the current study were less consistent in that both cue type and interferer condition were randomized from run to run.

Values of BII for some listeners fell quite close to zero, indicating little to no binaural interference. Some of those cases reflect realistic (non-ceiling) discrimination thresholds for interference and non-interference conditions (points near the diagonal in Fig. 2). Several cases, however, reflect poor baseline performance that did not change much when the interferer was added. For example, all three participants exhibiting BII near 0 for ITD condition 0R [0510, 1013, and 1308 in Fig. 3(b)] performed consistently near ceiling when the interferer was present. Individual differences of this type suggest that BII substantially underestimated the degree of interference for many listeners who could not do the task at a criterion level when the interferer was present. Future studies could investigate a somewhat easier task (e.g., using a longer stimulus) to tease apart such differences.

F. Apparent immunity from binaural interference in “transposed tones”

When a high-frequency carrier is modulated by a half-wave rectified low-frequency sinusoid, the resulting waveform exhibits several features similar to those of Gabor click trains: For example, the temporal envelope is comprised of a series of pulses with relatively steep rise/fall slopes. Also, pronounced off-periods occur between these pulses of sound. When parameters are chosen appropriately, the waveforms and spectra of the two stimulus types match very closely. For Gabor click trains, the pulse shape is Gaussian, overall bandwidth is controlled by pulse duration, and the off-period is controlled by the ICI. For sounds with half-wave rectified envelopes (transposed tones; [van de Par and Kohlrausch, 1997](#)), the pulse duration and the off period are each equal to half the modulation period (the ICI), and the pulse shape matches one half cycle of a sinusoid. The bandwidths of transposed tones are further controlled by low-pass filtering the rectified envelope ([Bernstein and Trahiotis, 2002](#)) to eliminate spectral components above 2000 Hz. The filtering step provides a close match to the low-pass Gaussian characteristic of the Gabor envelope used in the current study.

Gabor click trains and transposed tones—along with “raised-sine” stimuli in which the envelope is some exponent of a sinusoid ([John et al., 2002](#))—belong to a family of “pulsatile” stimuli characterized by off periods and steeply sloping pulse envelopes. All three of those stimuli have been shown to provide potent (and comparable) envelope ITD cues ([Haftner and Buell, 1990](#); [Bernstein and Trahiotis, 2002](#), [2009](#)). That potency presumably reflects the critical roles of off periods and envelope attack slopes for envelope ITD sensitivity in high-frequency amplitude modulation sounds ([Haftner, 1977](#); [Klein-Hennig et al., 2011](#)). As mentioned in the Introduction, ITD sensitivity also varies with envelope rate: $ICI > 10\text{--}12\text{ ms}$ provides optimal sensitivity to ongoing envelope ITD, whereas $ICI < 2\text{ ms}$ results in poor ongoing ITD sensitivity and strong dominance of the overall onset ([Haftner and Buell, 1990](#); [Bernstein and Trahiotis, 2002](#); [Stecker and Brown, 2010](#); [Stecker, 2014](#)). Moderate rates ($5\text{--}10\text{ ms ICI}$) provide intermediate sensitivity to the onset and ongoing cues.

[Bernstein and Trahiotis \(2004\)](#) reported that ITD targets carried by high-frequency (4 kHz) transposed tones were apparently immune to the effects of binaural interference. Stimuli in that study were presented at moderate rates (128 Hz; roughly 8 ms ICI). The result is consistent with the overall potency of the ITD cue for such stimuli. The current results suggest that higher-rate pulsatile stimuli (2-ms ICI or 500 Hz) are not immune to binaural interference, possibly because the ongoing cue is less potent at high rates. The observed dependence of binaural interference on target configuration (OR vs R0) suggests that the still-potent ITD of the overall onset—while not “immune”—is affected less than are the ongoing cues in such stimuli.

Two additional factors should be considered in relating these studies: First, the interferer used by [Bernstein and Trahiotis \(2004\)](#) was a 400-Hz band of noise centered at 500 Hz, whereas the current study employed a tonal interferer

matched to the target’s fundamental frequency. Thus, listeners may have experienced stronger fusion of the target and interferer in the current study (see Sec. III G). Indeed, several subjects reported great difficulty in “hearing out” the target when the interferer was present. Second, the stimulus duration was 32 ms in the current study vs 300 ms in that of [Bernstein and Trahiotis \(2004\)](#). The different durations, paired with greater access to ongoing ITD at moderate pulse rates, could have affected interference by altering the relative weight of onset and ongoing ITD cues (cf. [Tobias and Schubert, 1959](#)).

G. Binaural interference as cue-weighting within auditory objects

Binaural interference is widely described as the obligatory combination of binaural information across frequency bands. Yet, the degree to which this combination is truly “obligatory” varies significantly across individual listeners ([Woods and Colburn, 1992](#); [Heller and Trahiotis, 1995](#); [Best et al., 2007](#)) and across stimulus configurations. Binaural interference can be dramatically reduced by disrupting the temporal or harmonic relationship between target and interferer ([Trahiotis and Bernstein, 1990](#); [Buell and Haftner, 1991](#); [Hill and Darwin, 1996](#)) and by sequential grouping cues ([Best et al., 2007](#)). Because many of these manipulations also impact the perceptual fusion (vs segregation) of target and interferer, many authors have suggested that binaural interference reflects the perceptual grouping of frequency components in auditory object formation ([Woods and Colburn, 1992](#); [Hill and Darwin, 1996](#); [Stellmack and Lutfi, 1996](#); [Best et al., 2007](#); [Croghan and Grantham, 2010](#)). Others have noted that because interference often persists—even if in a reduced form—despite segregated perception, grouping *per se* is not a necessary aspect of binaural interference ([Buell and Trahiotis, 1993](#)).

It may be useful to consider the fusion of binaural information as potentially separate from other aspects of perceptual grouping (e.g., hearing out the pitch of a mistuned partial). For example, our current understanding of binaural cue weighting over time suggests that for slowly repeating sounds (e.g., 100-Hz click trains), each repetition provides a nearly independent sample of binaural information. Although they are not “objects” in the broader perceptual sense, the individual events can contribute independently to *binaural* perception. When their cue values agree, they can provide optimal summation ([Haftner and Dye, 1983](#)). When they disagree, listeners may perceive multiple intracranial locations even for a single auditory object. At higher rates (or for pure tones; see [Stecker and Bibee, 2014](#)), the events no longer contribute independently. Instead, the repetitions appear binaurally fused and the overall sound is localized on the basis of the most reliable cues (e.g., ITD at onset).

Applying similar logic to binaural interference, stimuli may be manipulated to favor or disfavor binaural fusion across frequency. Many of those manipulations appear to be the same as for cross-frequency grouping more generally (e.g., simultaneity and harmonicity), although binaural fusion might be neither a cause nor a consequence of object formation. Thus, as for temporal weighting, cross-frequency

weighting reveals a dissociation between fusion/segregation of binaural (“where”) and object (“what”) information: Binaural interference may be experienced even when a target tone is easily heard out (Stellmack and Dye, 1993), or a single grouped object may be experienced simultaneously in two locations (Haftner and Jeffress, 1968).

In the current study, we observed significant binaural interference (i.e., dominance of low-frequency ITD cues), onset dominance (greater weighting of ITD available at sound onset than offset), and a cumulative effect of the two (greater interference for sounds that lack ITD at onset). We believe that it is useful to consider these effects of temporal and spectral weighting in a unified framework, i.e., as a perceptual competition between multiple cues carried by various spectrotemporal components. That competition may be sensitive to grouping cues that enhance both the perception of single auditory objects and the spectrotemporal fusion of binaural cues.

IV. SUMMARY AND CONCLUSIONS

- (1) High-rate Gabor click trains are susceptible to binaural interference. Significant interference was observed for 500-Hz Gabor click trains across all conditions tested in the current study, in contrast to the apparent immunity of 128 Hz transposed tones to interference reported by Bernstein and Trahiotis (2004).
- (2) Greater binaural interference for ITD targets that lack onset cues (condition 0R). Targets with onset ITD (conditions R0 and RR) exhibited roughly twofold threshold elevations, consistent with previous literature on binaural interference using noise bands or sinusoidally amplitude-modulated tones. When ITD cues were only available after the onset in condition 0R, however, thresholds were elevated by a factor of eight or more. That is, the ongoing ITD cue appeared highly susceptible to binaural interference.
- (3) Similar interference effects for ILD targets with and without onset cues. Although significant interference effects were consistently observed for ILD targets, they were similar across conditions with (RR) and without (0R) onset cues. Of the three tested conditions, the greatest amount of binaural interference was, in fact, observed under condition R0, suggesting that late-arriving ILD cues contribute significantly to lateral discrimination even under conditions of binaural interference.
- (4) Large individual variation in binaural interference. As in past studies, we observed a wide range of binaural interference across listeners, suggesting variation in analytic listening and segregation of binaural information across frequency (Woods and Colburn, 1992; Stellmack and Dye, 1993).
- (5) Overall, results suggest that robust cues are not immune to binaural interference, but weak cues appear remarkably susceptible. That conclusion is, in our view, consistent with variance-weighted models of binaural cue integration (Buell and Haftner, 1991; Heller and Richards, 2010) and with the general notion that spectrotemporal weighting of binaural information is fundamentally similar to perceptual weighting in other (e.g., multisensory) domains.

ACKNOWLEDGMENTS

This research was supported by Grant No. R01 DC011548 (to G.C.S.) from the National Institute on Deafness and Other Communication Disorders (NIDCD). The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIDCD or the National Institutes of Health. Portions of this work were completed in conjunction with the first author’s AuD capstone research project, and were presented in poster form to the American Auditory Society (Scottsdale, AZ, March 2014). We thank Anna Diedesch for research contributions, Julie Bierer for excellent mentoring, and Brian Moore for his suggestion to quantify ITD thresholds using psychometric fits when adaptive efforts failed to capture the nature of the data. Anna Diedesch, Nate Higgins, Les Bernstein, and two anonymous reviewers provided helpful comments on earlier versions of the manuscript.

APPENDIX: ANALYTICAL COMPARISON OF PULSATILE STIMULI

Section III F of this manuscript compares the current results to those of Bernstein and Trahiotis (2004), who described an apparent immunity of transposed tones to binaural interference. In this manuscript, we argue that the essential similarity of transposed tones, Gabor click trains, and raised-sine stimuli (John *et al.*, 2002) justifies their treatment as a broader family of stimuli characterized by discrete off periods and steeply sloping envelopes. We refer to such stimuli as “pulsatile.”

Figure 4 presents a comparison of pulsatile and sinusoidally modulated stimuli. It is readily apparent that the waveforms and spectra of the various pulsatile stimuli agree quite closely when parameters are appropriately chosen. Figure 4(a) illustrates waveforms for carrier frequency of 4000 Hz and modulation rate of 500 Hz (2-ms ICI). Gabor pulse duration is set to $\sigma = 221 \mu\text{s}$, a typical value employed in the current study and in many past studies (Buell and Haftner, 1988; Haftner and Buell, 1990; Stecker and Haftner, 2002; Stecker and Brown, 2010; Stecker, 2014). For these parameters, off periods and envelope slopes match closely across the three types of pulsatile stimuli. Figure 4(b) compares the corresponding amplitude spectra, which are also similar. For all three stimuli, -40 dB bandwidth spans $\pm 2000 \text{ Hz}$ of the carrier frequency. The most notable difference is the omission of the $\pm 1.5 \text{ kHz}$ components in the transposed-tone spectrum, a consequence of even-order distortion induced by half-wave rectification of the sinusoidal envelope. Figure 4(c) illustrates the correspondence between envelope shapes at this modulation rate, which is particularly close between the Gabor click train (GCT) and the raised-sine stimulus with an exponent of four (RS4). Finally, Figure 4(d) illustrates temporal waveforms for a slower modulation rate (200 Hz or 5-ms ICI). Note that for SAM, transposed tones, and raised-sine stimuli, the pulse duration scales with the modulation period. For Gabor click trains, the two parameters are independent; effects of ICI have typically been studied by changing the off period without changing the shape of individual pulses (e.g., Stecker and Brown, 2010).

FIG. 4. Comparison of “pulsatile” stimuli: transposed tones (TT), Gabor click trains (GCT), raised-sine stimuli (RS4), and sinusoidally amplitude modulated (SAM) tones. (a) Temporal waveforms of all four stimuli. The carrier frequency is 4000 Hz in each case, and carrier phases have been adjusted to ensure alignment of peaks in the carrier and modulator waveforms. The modulation rate is 500 Hz (2-ms ICI). The raised-sine stimulus was generated with an exponent of four (Bernstein and Trahiotis, 2009). (b) Amplitude spectra of the transposed tone, Gabor click train, and raised-sine stimuli plotted in (a). Dotted horizontal lines indicate 20 dB range on the vertical axis. (c) Envelope of a single modulation cycle for each of the stimuli depicted in (b). (d) Temporal waveforms as in (a), but with modulation rate of 200 Hz (5 ms ICI). Note that GCT pulse duration remains fixed as in (a), while others scale with the modulation period.

Not surprisingly given their physical similarity, the various pulsatile stimuli appear to support similarly potent ITD cues. Table I lists the range of ITD thresholds reported by several studies that employed Gabor click trains, transposed tones, or raised-sine stimuli. Data using analog-filtered impulse trains (Hafta and Dye, 1983) are also included for comparison. Particularly given the range of performance variation across listeners, the clear overlap across studies suggests that stimulus differences are probably not meaningful after equating for the major variables of off period and attack slope (Klein-Hennig *et al.*, 2011; see also the discussion of graded changes in envelope peakedness/sharpness across raised-sine exponent by Bernstein and Trahiotis, 2009).

Before concluding, it should be noted that although the stimuli themselves can be made to match quite closely, differences in parameter selection across studies can make direct comparisons difficult. For example, studies using Gabor click trains have typically fixed the number of clicks and allowed overall duration to vary with ICI, whereas

studies of SAM and transposed tones have typically kept duration constant and allowed the number of modulation periods to vary. Also, Bernstein and Trahiotis (2002, 2009) and Monaghan *et al.* (2015) used 20-ms diotic onset ramps in an attempt to disrupt onset artifacts—e.g., ITD cues conveyed by spectral splatter and/or the onset envelope itself—whereas other studies did not. Bearing such differences in mind, we argue that comparisons of results across “different” pulsatile stimuli are both sensible and informative. In particular, we argue against the notion that any as-yet-unspecified features render specific stimuli uniquely capable of supporting performance or thwarting binaural interference. Rather, it is necessary to consider the range of parameters over which such effects appear. With respect to the topic of the current study, for example, “immunity” to binaural interference is not a specific feature of transposed tones, but more likely a general feature of pulsatile stimulation at moderate rates. Sensitivity to ongoing envelope ITD cues, in excess of that supported by SAM tones, has been repeatedly

TABLE I. Comparison of ITD thresholds across studies employing pulsatile stimulation.

Study	Stimulus	Threshold range
Buell and Hafter (1988)	12 Gabor clicks (4 ms ICI)	20–70 μ s
Bernstein and Trahiotis (2002)	300-ms transposed tone (256 Hz)	75–140 μ s
Bernstein and Trahiotis (2009)	300-ms raise-sine, exponent 4 (256 Hz)	100–250 μ s
Stecker and Brown (2010)	16 Gabor clicks (5 ms ICI)	40–170 μ s
Monaghan <i>et al.</i> (2015)	300-ms transposed tone (256 Hz)	40–400 μ s
Hafter and Dye (1983)	16 analog-filtered clicks (5 ms ICI)	30–80 μ s

demonstrated for such stimuli. Thus, the partial immunity of moderate-rate pulsatile stimuli to binaural interference can likely be understood in terms of the potency of available cues, as suggested by variance-weighted models of binaural cue integration (Buell and Hafter, 1991; Heller and Richards, 2010).

- Bernstein, L. R., and Trahiotis, C. (2002). “Enhancing sensitivity to interaural delays at high frequencies using ‘transposed stimuli,’” *J. Acoust. Soc. Am.* **112**, 1026–1036.
- Bernstein, L. R., and Trahiotis, C. (2004). “The apparent immunity of high-frequency ‘transposed’ stimuli to low-frequency binaural interference,” *J. Acoust. Soc. Am.* **116**, 3062–3069.
- Bernstein, L. R., and Trahiotis, 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.
- Best, V., Gallun, F. J., Carlile, S., and Shinn-Cunningham, B. G. (2007). “Binaural interference and auditory grouping,” *J. Acoust. Soc. Am.* **121**, 1070–1076.
- Brainard, D. H. (1997). “The psychophysics toolbox,” *Spat. Vis.* **10**, 433–436.
- Buell, T. N., and Hafter, E. R. (1988). “Discrimination of interaural differences of time in the envelopes of high-frequency signals: Integration times,” *J. Acoust. Soc. Am.* **84**, 2063–2066.
- Buell, T. N., and Hafter, E. R. (1991). “Combination of binaural information across frequency bands,” *J. Acoust. Soc. Am.* **90**, 1894–1900.
- Buell, T. N., and Trahiotis, C. (1993). “Interaural temporal discrimination using two sinusoidally amplitude-modulated, high-frequency tones: Conditions of summation and interference,” *J. Acoust. Soc. Am.* **93**, 480–4807.
- Croghan, N. B. H., and Grantham, D. W. (2010). “Binaural interference in the free field,” *J. Acoust. Soc. Am.* **127**, 3085–3091.
- David, E., Guttman, N., and Bergeijk, W. V. (1959). “Binaural interaction of high-frequency complex stimuli,” *J. Acoust. Soc. Am.* **31**, 774–782.
- Diedesch, A. C., and Stecker, G. C. (2015). “Temporal weighting of binaural information at low frequencies: Discrimination of dynamic interaural time and level differences,” *J. Acoust. Soc. Am.* **138**, 125–133.
- Efron, B., and Tibshirani, R. (1986). “Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy,” *Statist. Sci.* **1**, 54–75.
- Fox, J. (2008). *Applied Regression Analysis and Generalized Linear Models* (Sage, Thousand Oaks, CA), Chap. 21, pp. 587–606.
- Freyman, R. L., Zurek, P. M., Balakrishnan, U., and Chiang, Y. C. (1997). “Onset dominance in lateralization,” *J. Acoust. Soc. Am.* **101**, 1649–1659.
- Hafter, E. R. (1977). “Lateralization model and the role of time-intensity tradings in binaural masking: Can the data be explained by a time-only hypothesis?,” *J. Acoust. Soc. Am.* **62**, 633–635.
- Hafter, E. R., and Buell, T. N. (1990). “Restarting the adapted binaural system,” *J. Acoust. Soc. Am.* **88**, 806–812.
- Hafter, E. R., and Dye, R. H. J. (1983). “Detection of interaural differences of time in trains of high-frequency clicks as a function of interclick interval and number,” *J. Acoust. Soc. Am.* **73**, 644–651.
- Hafter, E. R., and Jeffress, L. A. (1968). “Two-image lateralization of tones and clicks,” *J. Acoust. Soc. Am.* **44**, 563–569.
- Harris, G. G. (1960). “Binaural interactions of impulsive stimuli and pure tones,” *J. Acoust. Soc. Am.* **32**, 685–692.
- Heller, L. M., and Richards, V. M. (2010). “Binaural interference in lateralization thresholds for interaural time and level differences,” *J. Acoust. Soc. Am.* **128**, 310–319.
- Heller, L. M., and Trahiotis, C. (1995). “Interference in detection of interaural delay in a sinusoidally amplitude-modulated tone produced by a second, spectrally remote sinusoidally amplitude-modulated tone,” *J. Acoust. Soc. Am.* **97**, 1808–1816.
- Hill, N. I., and Darwin, C. J. (1996). “Lateralization of a perturbed harmonic: Effects of onset asynchrony and mistuning,” *J. Acoust. Soc. Am.* **100**, 2352–2364.
- John, M. S., Dimitrijevic, A., and Picton, T. W. (2002). “Auditory steady-state responses to exponential modulation envelopes,” *Ear Hear.* **23**, 106–117.
- Klein-Hennig, M., Dietz, M., Hohmann, V., and Ewert, S. D. (2011). “The influence of different segments of the ongoing envelope on sensitivity to interaural time delays,” *J. Acoust. Soc. Am.* **129**, 3856–3872.
- McFadden, D., and Pasanen, E. G. (1976). “Lateralization of high frequencies based on interaural time differences,” *J. Acoust. Soc. Am.* **59**, 634–639.
- Monaghan, J. J. M., Bleack, S., and McAlpine, D. (2015). “Sensitivity to envelope interaural time differences at high modulation rates,” *Trends Hear.* **19**, 2331216515619331.
- Saberi, K. (1996). “Observer weighting of interaural delays in filtered impulses,” *Percept. Psychophys.* **58**, 1037–1046.
- Stecker, G. C. (2010). “Trading of interaural differences in high-rate Gabor click trains,” *Hear. Res.* **268**, 202–212.
- Stecker, G. C. (2014). “Temporal weighting functions for interaural time and level differences. IV. Effects of carrier frequency,” *J. Acoust. Soc. Am.* **136**, 3221–3232.
- Stecker, G. C., and Bibee, J. M. (2014). “Nonuniform temporal weighting of interaural time differences in 500 Hz tones,” *J. Acoust. Soc. Am.* **135**, 3541–3547.
- Stecker, G. C., and Brown, A. D. (2010). “Temporal weighting of binaural cues revealed by detection of dynamic interaural differences in high-rate Gabor click trains,” *J. Acoust. Soc. Am.* **127**, 3092–3103.
- Stecker, G. C., and Brown, A. D. (2012). “Onset- and offset-specific effects in interaural level difference discrimination,” *J. Acoust. Soc. Am.* **132**, 1573–1580.
- Stecker, G. C., and Gallun, F. J. (2012). “Binaural hearing, sound localization, and spatial hearing,” in *Translational Perspectives in Auditory Neuroscience: Normal Aspects of Hearing*, edited by K. L. Tremblay and R. F. Burkard (Plural, San Diego CA), Chap. 14, pp. 387–437.
- Stecker, G. C., and Hafter, E. R. (2002). “Temporal weighting in sound localization,” *J. Acoust. Soc. Am.* **112**, 1046–1057.
- Stellmack, M. A., and Dye, R. H., Jr. (1993). “The combination of interaural information across frequencies: The effects of number and spacing of components, onset asynchrony, and harmonicity,” *J. Acoust. Soc. Am.* **93**, 2933–2947.
- Stellmack, M. A., and Lutfi, R. A. (1996). “Observer weighting of concurrent binaural information,” *J. Acoust. Soc. Am.* **99**, 579–587.
- Tobias, J. V., and Schubert, E. R. (1959). “Effective onset duration of auditory stimuli,” *J. Acoust. Soc. Am.* **31**, 1595–1605.
- Trahiotis, C., and Bernstein, L. R. (1990). “Detectability of interaural delays over select spectral regions: Effects of flanking noise,” *J. Acoust. Soc. Am.* **87**, 810–813.
- van de Par, S., and Kohlrausch, A. (1997). “A new approach to comparing binaural masking level differences at low and high frequencies,” *J. Acoust. Soc. Am.* **101**, 1671–1680.
- Woods, W. S., and Colburn, H. S. (1992). “Test of a model of auditory object formation using intensity and interaural time difference discrimination,” *J. Acoust. Soc. Am.* **91**, 2894–2902.