Temporal weighting of binaural information at low frequencies: Discrimination of dynamic interaural time and level differences

Anna C. Diedesch and G. Christopher Stecker^{a)}

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

(Received 17 December 2014; revised 15 May 2015; accepted 27 May 2015; published online 8 July 2015)

The importance of sound onsets in binaural hearing has been addressed in many studies, particularly at high frequencies, where the onset of the envelope may carry much of the useful binaural information. Some studies suggest that sound onsets might play a similar role in the processing of binaural cues [e.g., fine-structure interaural time differences (ITD)] at low frequencies. This study measured listeners' sensitivity to ITD and interaural level differences (ILD) present in early (i.e., onset) and late parts of 80-ms pure tones of 250-, 500-, and 1000-Hz frequency. Following previous studies, tones carried static interaural cues or dynamic cues that peaked at sound onset and diminished to zero at sound offset or vice versa. Although better thresholds were observed in static than dynamic conditions overall, ITD discrimination was especially impaired, regardless of frequency, when cues were not available at sound onset. Results for ILD followed a similar pattern at 1000 Hz; at lower frequencies, ILD thresholds did not differ significantly between dynamic-cue conditions. The results support the "onset" hypothesis of Houtgast and Plomp [(1968). *J. Acoust. Soc. Am.* **44**, 807–812] for ITD discrimination, but not necessarily ILD discrimination, in low-frequency pure tones. © 2015 Acoustical Society of America. [<http://dx.doi.org/10.1121/1.4922327>]

[VB]

Pages: 125–133

I. INTRODUCTION

Auditory spatial information is conveyed by many different acoustic features (i.e., spatial cues) of sounds. Ideally, listeners would make use of, and appropriately weight, all available cues. In real rooms, for example, sound onsets carry reliable spatial cues while later sound can be distorted by echoes and reverberation. In that case, an appropriate strategy is to weight the spatial cues in a temporally nonuniform manner that emphasizes the early arriving sound. In fact, the auditory system does exactly this, as has been thoroughly established in the literature on the Franssen effect (Franssen, 1960; Hartmann and Rakerd, 1989; Yost *et al.*, 1997) and precedence effects (Wallach *et al.*, 1949; Yost and Soderquist, 1984; Zurek, 1987; Litovsky *et al.*, 1999; Brown *et al.*, 2015). When sounds carry spatial cues that are constant over time, however, emphasis of the onset cues provides no advantage; instead, listeners ought to benefit from weighting spatial cues in a temporally uniform manner. The results of several studies, however, suggest otherwise. Most recently, Stecker and Bibee (2014) reported nonuniform temporal weighting of static interaural time differences (ITD) in 500 Hz pure tones presented over headphones. That result was consistent with previous reports of nonuniform temporal weighting of ITD in bands of low-frequency noise (Houtgast and Plomp, 1968) and modulated high-frequency sounds (Hafters and Dye, 1983). In all three studies, ITD thresholds were measured as a function of sound duration. Despite differences in

the sounds employed, all three studies found that ITD thresholds improved with duration, but at a shallower rate than would be expected if listeners used all portions of the sound equally. This finding lead Houtgast and Plomp (1968, p. 811)—later echoed by Hafters and Dye (1983)—to conclude, “the onset of the signal contributes much more to [the accuracy of] the lateral position perceived than the ongoing part.”

Stecker and Brown (2010) and Stecker and Bibee (2014) tested that hypothesis more directly by comparing lateralization thresholds for sounds that differed in the availability of ITD at sound onset. In their approach, sounds carried ITD that was either constant throughout the duration (condition “RR”) or changed over time so that either the onset (in condition “OR”) or offset (“R0”) was diotic. Stecker and Brown (2010) presented high-rate trains of filtered clicks centered at 4000 Hz, similar to stimuli employed by Hafters and Dye (1983);¹ Stecker and Bibee (2014) adapted these methods to test pure tones at 500 Hz, the same frequency region tested by Houtgast and Plomp (1968). Both Stecker and Brown (2010) and Stecker and Bibee (2014) reported significant threshold elevations in condition OR (zero ITD at onset) compared to conditions in which the ITD cue was available near sound onset. That result parallels a recent observation by Dietz *et al.* (2013) of greater weighting of fine-structure ITD during the rising envelope (i.e., the “onset”), rather than the peak, of each modulation period in a sinusoidally amplitude-modulated 500 Hz tone. Here, we adopt the approach of Stecker and Brown (2010) to measure temporal weighting across a range of pure-tone frequencies for both ITD and for interaural level differences (ILD).

^{a)}Electronic mail: g.christopher.stecker@vanderbilt.edu

It is noteworthy that Houtgast and Plomp (1968), Hafter and Dye (1983), and Stecker and Bibee (2014) all found remarkably similar results despite the wide range of stimulus types (noises, click trains, and tones, respectively) and frequencies (500–4000 Hz) tested. As mentioned above, all three calculated improvements in lateralization accuracy with signal duration, and expressed these in the form of $\log(\text{threshold})$ -vs- $\log(\text{duration})$ slopes. For uniform temporal weighting, that slope should be -0.5 , corresponding to $1/\sqrt{(\text{duration})}$. Obtained values were significantly shallower but quantitatively similar across studies, ranging -0.23 to -0.25 for Houtgast and Plomp's two listeners, -0.08 to -0.33 (mean -0.22) for Hafter and Dye (1983), and -0.09 to -0.27 (mean -0.18) for Stecker and Bibee (2014). These remarkable similarities suggest that a common mechanism underlies temporal weighting of ITD, regardless of frequency. The literature on precedence effects, on the other hand, does suggest some differences across frequency. One important consideration is the interaction between successive stimuli on the basilar membrane (i.e., “ringing” of peripheral filters, see Tollin, 1998; Stecker, 2014). Because the temporal properties of the auditory peripheral response vary with cochlear place, one might expect larger effects at lower frequencies due to longer ringing. That expectation is roughly consistent with stronger precedence effects for stimuli with low-frequency early components and high-frequency late components than vice versa (Divenyi, 1992; Shinn-Cunningham *et al.*, 1995), and with the observations from Whitmer (2004) of stronger Franssen effects for low-frequency (250–500 Hz) than high-frequency (2000–4000 Hz) tones. Some studies have suggested additional effects of frequency, for example, stronger Franssen effects for mid-frequency (1000–1500 Hz) than higher- or lower-frequency tones (Yost *et al.*, 1997). To investigate potential differences across frequency, experiments of the current study investigated pure tones varying from 250 to 1000 Hz.

In contrast to the remarkable similarities in temporal weighting of ITD across studies, results regarding the temporal weighting of ILD have been more equivocal. Some studies have reported similar onset dominance for ILD and ITD (Zurek, 1980; Hafter *et al.*, 1983); others have suggested key differences, e.g., stronger precedence effects for ITD than for ILD cues (Krumbholz and Nobbe, 2002; Saberi *et al.*, 2004; Brown and Stecker, 2013). Hafter *et al.* (1983)

measured ILD thresholds as a function of duration and showed identical effects for ILD as they did for ITD (Hafter and Dye, 1983). However, Stecker and Brown (2010) found no evidence for specific onset weighting in ILD, despite using stimuli (4000-Hz click trains) similar to those of Hafter *et al.* (1983). Results of that and a follow-up study which compared the weighting of ILD carried by onsets, offsets, and interior portions of sounds (Stecker and Brown, 2012) suggest that ILD cues near sound offset are strongly weighted, unlike the case for ITD, where the onset dominates more completely. Here, we investigate whether similar differences between ITD and ILD exist for low-frequency pure tones.

II. EXPERIMENT 1: DISCRIMINATION OF DYNAMIC INTERAURAL TIME DIFFERENCES

A. Methods

Data collection took place in the Department of Speech and Hearing Sciences at the University of Washington. All procedures, including recruitment, consenting, and testing of human subjects, were in accordance to the guidelines of the University of Washington Human Subjects Division and were reviewed and approved by the cognizant Institutional Review Board.

1. Participants

Ten listeners, eight female, aged 20 to 40 (mean 26.4, SD 6.7) participated in this experiment. Four of the listeners had participated in a similar experiment (Stecker and Bibee, 2014). One of those was employed in the lab (listener 0510). Another of the listeners was the first author (0507), who did not participate in the previous study. Other listeners were native to the purpose of the study. A hearing screening conducted prior to data collection confirmed normal pure tone thresholds (<15 dB hearing level) at octave frequencies between 250 and 8000 Hz for all listeners. Participants other than lab personnel were monetarily compensated for their time.

2. Stimuli

Stimuli were pure tones of 80 ms duration presented at 60 dB sound pressure level (SPL) over closed circumaural earphones (Stax 4070). Tone frequency was 250, 500, or 1000 Hz, sound levels were calibrated to 60 dB SPL at each

FIG. 1. (Color online) Illustration of stimuli employed in experiment 1 (not to scale). In both experiments, listeners identified target intervals containing 80-ms pure tones with diotic envelopes and non-zero ITD (experiment 1) or ILD (experiment 2). Panels illustrate waveforms at each ear (line shading) with ITD imposed. Binaural differences were presented in three conditions: (left) The “RR” condition presented a constant ITD or ILD cue lateralized toward the right ear. In the dynamic-cue conditions “R0” and “0R,” ITD or ILD changed over time. (Center) In condition “R0,” the cue was largest at sound onset and diminished over time, while (right) in condition “0R,” the cue was largest at sound offset. Note that the waveforms in dynamic-cue conditions were diotic at sound offset (in condition R0) or sound onset (in condition 0R). To the extent that listeners rely primarily on binaural information present early in the sound, binaural-cue thresholds should be elevated in condition 0R relative to other conditions.

frequency using a head and torso simulator (Brüel & Kjær 4100D, Nærum, Denmark). Stimuli were gated on and off by diotic raised cosine ramps of 20-ms duration to minimize envelope cues (see Fig. 1). Reference stimuli were completely diotic (ITD = 0) whereas target stimuli carried a right-leading fine-structure ITD cue, Δt . ITD was applied in one of three configurations. In condition RR, ITD was static and equal to Δt throughout the duration of the tone. In condition R0, the ITD changed over time, diminishing linearly from Δt at sound onset to 0 μs at sound offset. The reverse was true in condition OR, in which ITD grew from 0 μs at sound onset to Δt at sound offset. Dynamic ITD values in these conditions were accomplished by introducing interaural frequency differences (<3.2 Hz, see below) while controlling the starting and ending phases in each ear. To reduce the reliability of that cue, overall frequency roved $\pm 10\%$ (e.g., ± 25 Hz at 250 Hz) between stimulus intervals. Intensity roved ± 5 dB.

3. Procedure

Participants were tested in a double-walled sound-attenuating chamber (IAC, New York, NY). At the beginning of each test block, a continuous diotic 500-Hz pure tone was presented over the headphones. Listeners were instructed to adjust the headphone placement until the sound appeared centered in head. Subsequently, experimental stimuli were presented in a four-interval, two-alternative forced-choice (4I2AFC) task, with the target occurring in either the second or third interval. Other intervals contained diotic reference tones. Responses to target stimuli were made on a four-button response box with light emitting diode (LED) lights located above each response button (TDT RBOX). LEDs blinked sequentially to mark the four intervals presented on each trial. On each trial, listeners pressed one of the buttons to indicate which interval contained the right-leading target. Feedback was provided to the listeners on

each trial. Throughout the experiment, participants were in control of the pace of testing and were given breaks at least every 30 min or as needed.

ITD thresholds for each test block were estimated using a two-down one-up adaptive procedure tracking Δt at 71% correct in two interleaved but independent tracks (Levitt, 1971). Initial values of Δt depended on frequency condition and were set to 600 μs , 500 μs , or 250 μs , at 250, 500, and 1000 Hz, respectively. Δt was adjusted by a scaling factor of 0.2 for the first 4 of 12 reversals and 0.05 thereafter. ITD threshold was estimated as the geometric mean of Δt during the final eight reversals recorded in each track.

Prior to the start of the randomized experimental conditions, participants were presented with written and verbal instructions, followed by at least four practice blocks (eight threshold measurements) in condition RR at 250, 500, and 1000 Hz. Additional verbal instructions, given after the first practice block, instructed listeners to attend only to the lateralization of the sounds and to ignore fluctuations in pitch and intensity. Experimental test blocks followed, each presenting one combination of the three conditions (RR, R0, and OR) and three test frequencies (250, 500, 1000 Hz). After each combination was tested once (in random order), the entire set of nine combinations was repeated three additional times, for a total of four test blocks (eight interleaved tracks) per combination. In cases where an adaptive track failed to behave asymptotically, the data were eliminated and the test block was repeated. Test blocks lasted approximately 8 to 10 min, and data were collected in 2-hour test sessions. Each listener completed approximately 8 to 10 such sessions.

4. Analysis

Data were analyzed following the approach of Stecker and Bibee (2014), who describe that approach in more detail.

FIG. 2. Individual ITD thresholds from experiment 1. In each panel, symbols plot ITD thresholds (vertical axis) for conditions RR (black squares), R0 (rightward-pointing triangles), and OR (leftward-pointing triangles) against tone frequency (horizontal axis). Error bars indicate bootstrapped 95% confidence intervals. Light gray lines plot two times the RR thresholds for comparison. Individual listeners' data are represented in separate panels. In most cases, ITD thresholds improved with increasing frequency from 250 to 1000 Hz, and with the availability of ITD information at sound onset (best in condition RR and worst in condition OR).

FIG. 3. Group-mean thresholds from experiment 1. Symbols plot group-mean normalized ITD thresholds (vertical axis) against frequency (horizontal axis) in conditions RR (squares), R0 (rightward-pointing triangles), and OR (leftward-pointing triangles). Means were computed following normalization to each listener's RR threshold at the corresponding frequency. Error bars represent bootstrapped 95% confidence intervals.

Briefly, geometric mean thresholds for individual listeners were computed across test blocks in each experimental condition (see Fig. 2). For group-level analyses, each such threshold was normalized via division by the threshold obtained in condition RR for the corresponding participant and frequency. Group-average normalized thresholds were then computed across listeners (Fig. 3). Bootstrapped 95% confidence intervals were computed by 1000-fold resampling of threshold estimates across test blocks for individual listeners (Efron and Tibshirani, 1986), and second-level resampling of those data across listeners for group data (Stecker and Bibee, 2014).

To test the “onset” hypothesis described by Houtgast and Plomp (1968) and Stecker and Bibee (2014), we computed ratios of threshold Δt obtained in conditions OR and R0 (Fig. 4). Bootstrapped 95% confidence intervals were computed by resampling the threshold estimates in each condition and computing the corresponding ratio for each bootstrapped sample. Similarly at the group level, confidence intervals were computed by second-level resampling of the group-average thresholds, prior to calculating the OR/R0 threshold ratio for each bootstrapped sample. The proportion of such ratio values falling at or below 1 (i.e., the p-value of a one-tailed confidence interval) quantified the statistical significance that OR thresholds exceeded R0 thresholds at the group level.

Finally, the effects of frequency were assessed using a factorial repeated measures analysis of variance (ANOVA) on normalized thresholds, with factors of participant, condition (RR, R0, OR), and frequency (250, 500, 1000 Hz).

FIG. 4. Bars plot the ratio of OR to R0 ITD thresholds obtained in experiment 1 for each individual (horizontal axis), and for the group average (far right). Bar shading indicates tone frequency. Error bars represent bootstrapped 95% confidence intervals. Consistent with Figs. 2 and 3, most cases indicate higher thresholds in condition OR than in condition R0 (i.e., ratio >1).

TABLE I. Group-mean threshold values, computed without normalization, for experiment 1 (ITD in μs , left) and experiment 2 (ILD in dB, right). Means and standard deviations (in parentheses) are given for each combination of frequency (columns) and condition (rows).

	Experiment 1: ITD Thresholds (μs)			Experiment 2: ILD Thresholds (dB)		
	250 Hz	500 Hz	1000 Hz	250 Hz	500 Hz	1000 Hz
RR	140.5 (69.6)	59.4 (27.0)	53.5 (23.0)	2.9 (1.5)	2.9 (1.4)	4.2 (2.0)
R0	178.9 (53.8)	84.1 (23.3)	79.2 (33.8)	4.9 (2.7)	4.6 (1.5)	5.7 (1.5)
OR	237.6 (71.7)	132.4 (44.8)	111.9 (39.1)	5.1 (2.5)	5.3 (1.9)	6.9 (2.1)

B. Results

Individual listeners' ITD thresholds for the nine stimulus conditions (symbols) are plotted against frequency in Fig. 2. Error bars indicate 95% bootstrapped confidence intervals. Gray lines in each panel plot double the threshold obtained in condition RR (squares), for comparison to dynamic-cue thresholds obtained in conditions R0 (rightward-pointing triangles) and OR (leftward-pointing triangles). Four general observations can be made from these data: First, individual listeners varied in overall sensitivity—with thresholds ranging from 29 to 105 μs in condition RR at 500 Hz, for example—but demonstrated similar patterns of threshold variation across conditions. Second, thresholds tended to decrease with increasing frequency. Third, in almost every case, the highest thresholds were obtained in condition OR. That result is consistent with the “onset” hypothesis, since that stimulus carries zero ITD at sound onset. Fourth, and less consistent with that hypothesis, better thresholds were obtained in condition RR than R0, suggesting that listeners did benefit from consistent ITD information late in the stimulus, when available. Frequency- and condition-dependent differences are further corroborated in Table I (experiment 1), which displays means and standard deviations of non-normalized threshold data across listeners.

Figure 3 plots group-average thresholds, normalized to condition RR at each frequency (squares). Triangles plot values for stimulus conditions R0 and OR, as in Fig. 2. Error bars represent bootstrapped 95% confidence intervals. Similarly to the individual data, higher thresholds were observed in condition OR than R0, both of which significantly exceeded RR thresholds (i.e., normalized thresholds >1). The overall effect of frequency (Fig. 2) cannot be seen, because data were normalized separately at each frequency.

Factorial repeated-measures ANOVA indicated no significant effects of frequency, or frequency-by-condition

interaction, on normalized threshold values in conditions R0 and 0R; $F(2,18) = 1.90$; $p = 0.18$, $F(2,18) = 1.06$; $p = 0.37$, respectively. Thus, it appears that the cross-frequency differences apparent in Fig. 2 were independent of stimulus-condition effects, and were removed by the frequency-specific threshold normalization. ANOVA did reveal a significant main effect of stimulus condition (R0, 0R), $F(1,9) = 89.51$; $p < 0.05$, consistent with the trends apparent in Figs. 2 and 3.

The “onset” hypothesis for ITD discrimination can be meaningfully addressed by directly comparing thresholds obtained in conditions R0 and 0R. For a given Δt , the long-term average ITD is equal across the two conditions, which differ only in the temporal configuration of the cue. According to the “onset” hypothesis, listeners should be significantly more sensitive to ITD in condition R0 than 0R. To compare these two conditions, the threshold ratio $\Delta t_{0R}/\Delta t_{R0}$ was calculated for each listener and for the group average, all of which are plotted in Fig. 4. Although ratios varied across conditions and listeners, they were consistently ≥ 1 . Average ratio values were greater than 1 overall: 1.3 ($p < 0.005$), 1.6 ($p < 0.001$), and 1.4 ($p < 0.001$) at 250, 500, and 1000 Hz, respectively. These results strongly support the “onset” hypothesis for ITD across this frequency range.

III. EXPERIMENT 2: DISCRIMINATION OF DYNAMIC INTERAURAL LEVEL DIFFERENCES

A. Methods

Aside from differences in the stimuli presented, which carried ILD rather than ITD, experimental methods corresponded closely to experiment 1. Except where mentioned specifically below, methodological details were identical to those described above.

1. Participants

The same ten listeners from experiment 1 participated in experiment 2.

2. Stimuli

Stimuli were identical to those of experiment 1 with the exception of interaural cue type, which was ILD in experiment 2. Reference stimuli were diotic (ILD = 0) whereas target stimuli carried an ILD cue, ΔL , which favored the right ear. As in experiment 1, ILD was applied in one of three configurations: RR, R0, and 0R. In condition RR, ILD was static and equal to ΔL throughout the duration of the tone. In condition R0, the ILD changed over time, diminishing linearly (in dB) from ΔL at sound onset to 0 dB at sound offset. The reverse was true in condition 0R, in which ILD grew from 0 dB at sound onset to ΔL at sound offset. Also as in experiment 1, overall frequency roved $\pm 10\%$ between intervals and intensity roved ± 5 dB.

3. Procedure

For ILD threshold measurement, adaptive tracks were initialized $\Delta L = 10$ dB for all test frequencies. ΔL was adjusted in steps of 2 dB for the first two reversals, 0.5 dB for the next four reversals, and 0.1 dB for the final eight reversals. The arithmetic mean of ΔL (in dB) on those last eight reversals was the estimated ILD threshold.

4. Analysis

Analysis of ILD data differed from ITD analyses only with respect to the units involved, i.e., differences in dB ILD versus ratios in μs ITD. Thus, arithmetic rather than geometric mean thresholds were computed across test blocks for individual listeners (see Fig. 5). Group-average thresholds were computed following normalization by subtraction in

FIG. 5. Individual ILD thresholds from experiment 2. In each panel, symbols plot ILD thresholds (vertical axis) for conditions RR (black squares), R0 (rightward-pointing triangles), and 0R (leftward-pointing triangles) against tone frequency (horizontal axis). Error bars indicate bootstrapped 95% confidence intervals. Individual listeners' data are represented in separate panels. Light gray lines plot two times the RR thresholds for comparison. Overall, ILD thresholds for dynamic cue conditions R0 and 0R were poorer than for condition RR by roughly this amount. Differences between R0 and 0R thresholds appeared smaller, for most listeners, than in experiment 1.

FIG. 6. Group-mean thresholds from experiment 2. Symbols plot group-mean ILD thresholds (vertical axis) against frequency (horizontal axis) in conditions RR (squares), R0 (rightward-pointing triangles), and OR (leftward-pointing triangles). Means were computed following normalization to each listener's RR threshold at the corresponding frequency. Error bars represent bootstrapped 95% confidence intervals. Mean ILD thresholds in the R0 condition improved with increasing frequency, whereas OR thresholds grew. Thus, ILD thresholds for OR were significantly poorer than for R0 at 500 Hz ($p < 0.05$) and 1000 Hz ($p < 0.01$). As in experiment 1, ILD thresholds obtained in both dynamic-cue conditions were poorer than in condition RR (normalized thresholds > 0 , $p < 0.01$).

dB, rather than division, of each participant's mean threshold obtained in condition RR from that participant's mean thresholds obtained in other conditions (Fig. 6). Similarly, to test the "onset" hypothesis for ILD, we computed the difference in thresholds ΔL obtained in condition OR and R0 (Fig. 7).

B. Results

Individual listeners' ILD thresholds for the nine stimulus conditions (symbols) are plotted against frequency in Fig. 5, which is formatted identically to Fig. 2. As in experiment 1, individual listeners varied in overall sensitivity. For example, in condition RR at 1000 Hz thresholds ranged from 2.0 to 7.2 dB. But overall, the majority of listeners demonstrated similar patterns of threshold variation across conditions (also evident in means of non-normalized data displayed in Table I, experiment 2). Specifically, the lowest thresholds were observed in condition RR, and thresholds in all three conditions tended to increase with increasing frequency for most listeners.² Differences between ILD thresholds obtained in conditions OR and R0 were not consistent across listeners. As compared to ITD thresholds obtained in experiment 1, fewer listeners showed ILD-threshold

elevations for condition OR relative to R0. Only one listener (1113), in fact, exhibited that pattern at all three frequencies.

Figure 6 plots group-average thresholds, normalized to condition RR at each frequency, formatted as in Fig. 3. Similarly to the individual data, thresholds observed in conditions OR and R0 significantly exceeded RR thresholds (i.e., normalized thresholds > 0). Thresholds in conditions OR and R0 did not differ at 250 Hz but diverged with increasing frequency, with worse thresholds in condition OR than R0 at 500 Hz ($p < 0.05$) and 1000 Hz ($p < 0.01$).

Factorial repeated-measures ANOVA indicated no significant main effect of frequency on normalized thresholds in conditions R0 and OR; $F(2,18) = 0.003$; $p = 0.997$. As for ITD thresholds in experiment 1, cross-frequency differences in overall threshold were thus apparently removed by the frequency-specific threshold normalization. ANOVA did reveal a significant main effect of stimulus condition (R0, OR), $F(1,9) = 8.04$; $p < 0.05$. The frequency-by-condition interaction showed a non-significant trend [$F(2,18) = 3.45$; $p = 0.054$] consistent with the data shown in Fig. 6.

Similarly to experiment 1, the "onset" hypothesis was tested by calculating the threshold difference $\Delta L_{OR} - \Delta L_{R0}$ for each listener and for the group average. These are plotted in Fig. 7. Threshold differences varied across conditions and listeners but were consistently ≥ 0 at 500 and 1000 Hz. Average difference values were 0.2 dB ($p = 0.281$), 0.7 dB ($p < 0.05$), and 1.2 dB ($p < 0.01$) at 250, 500, and 1000 Hz, respectively. These results only weakly support the "onset" hypothesis for ILD and only at 500 and 1000 Hz.

IV. DISCUSSION

A. Comparison of binaural sensitivity at sound onsets and offsets

The current study extends the work of Stecker and Bibee (2014) to evaluate the temporal weighting of ILD as well as ITD across a broader range of pure-tone frequencies than assessed in that study. The results for ITD at 500 Hz are remarkably consistent with those of the previous study, indicating ~ 1.5 -fold threshold elevation in condition OR relative to R0. Both studies share the methodological approach first used by Stecker and Brown (2010) with high-frequency click trains. For ITD, similar results were consistently observed in all three studies regardless of the stimulus type and carrier frequency. The static cue condition (RR) consistently produced the best ITD thresholds and the OR condition the

FIG. 7. Bars plot the differences between ILD thresholds obtained in conditions OR and R0 of experiment 2, for each listener (horizontal axis) and for the group average (far right). Bar shading indicates tone frequency. Error bars represent bootstrapped 95% confidence intervals. On average, normalized ILD thresholds were similar between conditions R0 and OR (i.e., difference ~ 0), but diverged at higher frequencies (i.e., difference > 0).

poorest, with performance on the R0 condition falling somewhere in between.

Consistent with the “onset” hypothesis (Houtgast and Plomp, 1968; Hafter and Dye, 1983; Hafter *et al.*, 1983; Stecker and Brown, 2010; Stecker and Bibee, 2014), the results suggest that listeners place greater weight on ITD cues appearing early in the sound (i.e., at sound onset). That is, removing the informative ITD cue from the early/onset portion in condition 0R consistently and significantly impaired thresholds in comparison to conditions with informative ITD at sound onset. The similarity of results across studies employing both envelope and fine-structure ITD, and stimuli ranging from low-frequency pure tones (the current study) to noise bands (Houtgast and Plomp, 1968) and click trains (Hafter and Dye, 1983), suggests a universal importance of ITD cues near sound onset.

The case for ILD is not as clear. Hafter and Dye (1983) and Hafter *et al.* (1983) found virtually identical results for high-frequency click trains carrying ITD or ILD, respectively, leading Hafter *et al.* (1983) to conclude that sound onsets play similarly important roles in the processing of both cues. In contrast, Stecker and Brown (2010), using very similar stimuli, found evidence of onset weighting for ITD but not for ILD. Our data for low-frequency pure tones appear consistent with Hafter *et al.* (1983) at 1000 Hz, but with Stecker and Brown (2010) at lower frequencies (250 Hz).

Stecker and Brown (2012) suggested that the discrepancies between these studies resulted from effects of onset and offset weighting in ILD processing (i.e., “U-shaped” temporal weighting functions, Stecker and Hafter, 2009). It is possible that similar issues pertain to the current data, although the stimuli and frequency range are quite different. Some support for that possibility is given by recent results demonstrating that some aspects of onset/offset weighting become more pronounced with decreasing carrier frequency (Stecker, 2014).

B. Frequency differences

As depicted in Figs. 2 and 5, as well as Table I, ITD and ILD thresholds varied overall as a function of sound frequency. Frequency-dependent effects were largely consistent across listeners but differed between ITD and ILD. Repeated-measures factorial ANOVA on non-normalized data confirmed the significant main effect of frequency on ITD thresholds [$F(2,18) = 41.58$, $p < 0.05$]. The tested frequencies were all within the range of accurate ITD coding in human listeners, and the trend of lower thresholds for higher frequency is consistent with previous data (Klumpp and Eady, 1956; Zwislocki and Feldman, 1956; Mills, 1958). Quantitatively, thresholds at 250 Hz were elevated by a factor of 2.6 relative to 1000 Hz, comparable to elevations of 2.3 (Zwislocki and Feldman, 1956) and 2.5 (Klumpp and Eady, 1956) reported previously, despite the longer duration and correspondingly lower thresholds obtained in those studies. Threshold differences across frequency might primarily reflect the relationship between ITD and interaural phase (i.e., for a given threshold phase difference, the ITD

equivalent is inversely proportional to the frequency). Another potentially relevant issue in the current study, and in previous studies, is that stimuli were presented at the same duration (here, 80 ms) across all frequencies. This corresponds to 80 cycles of the 1000-Hz tone, but only 20 cycles at 250 Hz. If thresholds are constrained by the number of available cycles, one might therefore expect better performance at higher frequencies, as the data show.

Repeated-measures factorial ANOVA also revealed a significant main effect of frequency on ILD thresholds [$F(2,18) = 4.85$, $p < 0.05$]. The direction of that effect was opposite to that for ITD. In general, ILD thresholds increased with frequency from 250 to 1000 Hz, consistent with previous data (Mills, 1960; Grantham, 1984) indicating an elevation of ILD thresholds around 1000 Hz. Despite large individual differences in overall thresholds, patterns of frequency dependence were remarkably similar across listeners (i.e., increasing with frequency for 8 of 10 listeners). Although the mechanism(s) responsible for these differences are not well understood, they could potentially relate to the divergence of 0R and R0 thresholds with frequency observed for ILD (see Fig. 6). That is, at least for some listeners, ILD thresholds in condition 0R appear to be more sensitive to frequency than ILD thresholds in condition R0.

C. Relating sound onsets to fluctuations in ongoing envelopes

The current results support a significant role for fine-structure ITD cues that co-occur with the onset of a brief but otherwise steady tone. In that regard, they compare favorably to recent work by Dietz *et al.* (2013), who demonstrated that listeners’ lateral perception of a periodically amplitude modulated 500-Hz tone was dominated by the fine-structure ITD cues which co-occurred with the rising flanks of the periodic envelope. As in the current study, stimuli presented slight interaural frequency differences, leading to dynamic ITD that drifted from one ear to the other over the course of each modulation cycle. To some extent, this “amplitude-modulated binaural beat” stimulus acted like a series of shorter dynamic-ITD tones as presented in the current study.³ In both cases, the overall amplitude envelope was diotic and the fine-structure ITD varied smoothly throughout the duration. In both studies, binaural perception was dominated by the fine-structure ITD cue present during the rising envelope fluctuation. It is possible that both results reflect a common mechanism that exploits envelope fluctuations to guide binaural cue extraction across a wide range of stimuli. If so, several connections can be made to other relevant aspects of the literature on the processing of ongoing binaural cues:

First, the ability to utilize these cues may depend on the rate of amplitude modulations. At high carrier frequency, modulations slower than ~ 100 Hz provide equal quanta of binaural information to be extracted from each modulation event (Hafter and Dye, 1983); more rapid modulations result in strong onset dominance consistent with binaural cue extraction only at the single overall onset. In that case, however, introducing a brief gap in the middle of the stimulus

provides another quantum of information, perhaps due to “re-triggering” of binaural cue extraction by the consequent envelope fluctuation (Hafter and Buell, 1990).

Second, highly fluctuating sounds such as noises should be capable of signaling binaural information in a continuous manner throughout the ongoing portion of the sound, due to the inherent fluctuations of the ongoing envelope. Indeed, numerous studies have demonstrated dominance of ongoing binaural cues over onset cues for noise targets (e.g., Tobias and Schubert, 1959; Freyman *et al.*, 1997), and greater sensitivity to ongoing cues for sounds with stochastic versus periodic temporal features (Goupell *et al.*, 2009; Brown and Stecker, 2011).

Third, and finally, temporally sparse stimuli with large fluctuations, such as speech, should provide strong envelope cues to guide robust binaural cue extraction in noisy and complex auditory scenes. This expectation suggests clear implications for the relationship between temporal envelope processing abilities (e.g., gap detection) and spatial hearing in normal and impaired listeners (Strouse *et al.*, 1998; Ochi *et al.*, 2014), and perhaps for the design of algorithms for binaural cue enhancement in assistive listening devices.

D. Individual differences

Apparent in both Fig. 4 and Fig. 7 are considerable individual differences in the degree of threshold elevation in condition 0R relative to R0 across listeners. Although no individuals show consistent trends in the opposite direction (R0 worse than 0R), several show negligible differences in conditions where others show large differences. An important question is whether individual listeners exhibit similar patterns of temporal weighting regardless of cue type (ITD or ILD). This was tested by correlating the 0R/R0 ITD threshold ratios (Fig. 4) to the 0R/R0 ILD threshold differences (Fig. 7) across listeners, separately at each tested frequency. Statistical significance of these correlations was assessed by 1000-fold bootstrapping across listeners (Efron and Tibshirani, 1986). In no case was that correlation statistically significant (250 Hz: $R^2 = 0.005$, $p = 0.8$; 500 Hz: $R^2 = 0.01$, $p = 0.8$; 1000 Hz: $R^2 = 0.02$, $p = 0.7$), suggesting that temporal weighting does not necessarily influence all cues in the same manner.

A second question is whether these effects correlate across frequency. Although static-cue thresholds in several conditions did correlate across adjacent frequencies (ITD, 250 vs 500 Hz: $R^2 = 0.53$, $p = 0.02$; 500 vs 1000 Hz: $R^2 = 0.31$, $p = 0.1$; ILD, 250 vs 500 Hz: $R^2 = 0.58$, $p = 0.01$; 500 vs 1000 Hz: $R^2 = 0.59$, $p = 0.01$), 0R/R0 threshold ratios did not (ITD, 250 vs 500 Hz: $R^2 = 0.29$, $p = 0.1$; 500 vs 1000 Hz: $R^2 = 0.10$, $p = 0.4$; ILD, 250 vs 500 Hz: $R^2 = 0.16$, $p = 0.3$; 500 vs 1000 Hz: $R^2 = 0.12$, $p = 0.3$). Thus, although listeners do appear consistent across frequency in their overall sensitivity to (static) ITD and ILD, temporal weighting may be more frequency-specific. Alternatively, the limited correlations of 0R/R0 ratio across frequency and across cue type could simply reflect greater listener-to-listener variability.

E. The physiological bases of these effects

The results suggest that in many cases, binaural cues conveyed early in the signal, perhaps those conveyed during the sound onset, exert a greater influence than do cues conveyed later in the sound. A number of potential mechanisms might contribute, including destructive interference on the basilar membrane (Tollin, 1998) or adaptation in the neuronal inputs to binaural comparison (as suggested by Hafter, 1997; Dietz *et al.*, 2014). More central mechanisms could also potentially contribute, such as interactions between envelope sensitivity and binaural tuning among neurons of the superior olivary complex (Remme *et al.*, 2014) and inferior colliculus (Mao and Carney, 2015).

V. SUMMARY AND CONCLUSIONS

- (1) When fine-structure ITD cues were available at sound onset, but not offset, discrimination thresholds were consistently better than in the opposite case (ITD available at sound offset but not onset). When the cue was available throughout the sound duration, thresholds were consistently lowest of all.
- (2) These differences in ITD discrimination were independent of frequency across the tested range of 250–1000 Hz
- (3) The results for ITD are quantitatively similar to previous studies using the same method to test 500 Hz pure tones (Stecker and Bibee, 2014) and rapid high-frequency click trains (Stecker and Brown, 2010).
- (4) Overall, these results support the “onset” hypothesis of Houtgast and Plomp (1968; see also Hafter and Dye, 1983). That is, discrimination of ITD—across a wide range of stimuli—relies heavily on the cues present in the earliest part of a sound.
- (5) A similar advantage of sound onsets was observed for ILD discrimination, but varied with tone frequency. Dynamic ILD discrimination thresholds significantly favored sound onsets at 1000 Hz, but not at 250–500 Hz. As for ITD, best performance was observed when the cue was constant with duration.
- (6) The pattern of results observed for ILD is thus less consistent with the “onset” hypothesis, although the general trend of the data appears to favor it except at the lowest frequencies tested.

ACKNOWLEDGMENTS

The authors thank Julie Stecker and Jackie Bibee for their assistance in recruiting participants, and Gin Best and two anonymous reviewers for helpful comments on earlier versions of the manuscript. Portions of this work were previously presented as a poster at the 164th meeting of the Acoustical Society of America in Kansas City, MO. This work was supported by NIH NIDCD [R01-DC011548].

¹Stecker and Brown (2010) and Hafter and Dye (1983) measured these effects as a function of inter-click interval (ICI) ranging from 1 to 10 ms. At the shortest values (i.e., highest rates), both studies found strong onset effects. For longer values (i.e., slower rates), temporal weighting was

relatively more uniform. In our view the high-rate/short-ICI data are most relevant to those of the current study; our comments reflect those data specifically.

²Note that, for several of the listeners, error bars appear systematically larger for high-threshold data points. This suggests that the psychophysical dimension related to ILD may not be linear in units of dB (i.e., psychometric functions not only shift, but become shallower, in more difficult conditions).

³Each period of the stimulus presented by Dietz *et al.* (2013) was 31.25 ms in duration, compared to the 80-ms duration employed here, and sinusoidal in shape. That is, it lacked the 40-ms steady portion of the tones presented here. The rise and fall characteristics were thus quite similar: here, we employed 20-ms $\cos^2$ ramps at onset and offset, compared to identically shaped 15.625-ms ramps in Dietz *et al.* (2013).

- Brown, A. D., and Stecker, G. C. (2011). "Temporal weighting functions for interaural time and level differences. II. The effect of binaurally synchronous temporal jitter," *J. Acoust. Soc. Am.* **129**, 293–300.
- Brown, A. D., and Stecker, G. C. (2013). "The precedence effect: Fusion and lateralization measures for headphone stimuli lateralized by interaural time and level differences," *J. Acoust. Soc. Am.* **133**, 2883–2898.
- Brown, A. D., Stecker, G. C., and Tollin, D. J. (2015). "The precedence effect in sound localization," *J. Assoc. Res. Otolaryngol.* **16**(1), 1–28.
- Dietz, M., Marquardt, T., Salminen, N. H., and McAlpine, D. (2013). "Emphasis of spatial cues in the temporal fine structure during the rising segments of amplitude-modulated sounds," *Proc. Natl. Acad. Sci.* **110**(37), 15151–15156.
- Dietz, M., Marquardt, T., Stange, A., Pecka, M., Grothe, B., and McAlpine, D. (2014). "Emphasis of spatial cues in the temporal fine structure during the rising segments of amplitude-modulated sounds II: Single-neuron recordings," *J. Neurophysiol.* **111**, 1973–1985.
- Divenyi, P. L. (1992). "Binaural suppression of nonechoes," *J. Acoust. Soc. Am.* **91**, 1078–1084.
- Efron, B., and Tibshirani, R. (1986). "Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy," *Stat. Sci.* **1**, 54–75.
- Franssen, N. V. (1960). "Some considerations on the mechanism of directional hearing," Ph.D. dissertation, Technische Hogeschool, Delft, the Netherlands.
- Freyman, R. L., Zurek, P. M., Balakrishnan, U., and Chiang, Y. C. (1997). "Onset dominance in lateralization," *J. Acoust. Soc. Am.* **101**, 1649–1659.
- Goupell, M. J., Laback, B., and Majdak, P. (2009). "Enhancing sensitivity to interaural time differences at high modulation rates by introducing temporal jitter," *J. Acoust. Soc. Am.* **126**, 2511–2521.
- Grantham, D. W. (1984). "Interaural intensity discrimination: Insensitivity at 1000 Hz," *J. Acoust. Soc. Am.* **75**, 1191–1194.
- Haftner, E. R. (1997). "Binaural adaptation and the effectiveness of a stimulus beyond its onset," in *Binaural and Spatial Hearing in Real and Virtual Environments*, edited by H. Gilkey and T. R. Anderson (Lawrence Erlbaum Associates, Mahwah, NJ), pp. 211–232.
- Haftner, E. R., and Buell, T. N. (1990). "Restarting the adapted binaural system," *J. Acoust. Soc. Am.* **88**, 806–812.
- Haftner, 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 number," *J. Acoust. Soc. Am.* **73**, 644–651.
- Haftner, E. R., Dye, R. H. J., and Wenzel, E. (1983). "Detection of interaural differences of intensity in trains of high-frequency clicks as a function of interclick interval and number," *J. Acoust. Soc. Am.* **73**, 1708–1713.
- Hartmann, W. M., and Rakerd, B. (1989). "Localization of sound in rooms IV: The Franssen effect," *J. Acoust. Soc. Am.* **86**, 1366–1373.
- Houtgast, T., and Plomp, R. (1968). "Lateralization threshold of a signal in noise," *J. Acoust. Soc. Am.* **44**, 807–812.
- Klumpp, R. G., and Eady, H. R. (1956). "Some measurements of interaural time difference thresholds," *J. Acoust. Soc. Am.* **28**, 859–860.
- Krumbholz, K., and Nobbe, A. (2002). "Buildup and breakdown of echo suppression for stimuli presented over headphones—The effects of interaural time and level differences," *J. Acoust. Soc. Am.* **112**, 654–663.
- Levitt, H. (1971). "Transformed up-down methods in psychoacoustics," *J. Acoust. Soc. Am.* **49**, 467–477.
- Litovsky, R. Y., Colburn, H. S., Yost, W. A., and Guzman, S. J. (1999). "The precedence effect," *J. Acoust. Soc. Am.* **106**, 1633–1654.
- Mao, J., and Carney, L. H. (2015). "Tone-in-noise detection using envelope cues: Comparison of signal-processing-based and physiological models," *J. Assoc. Res. Otolaryngol.* **16**(1), 121–133.
- Mills, A. W. (1958). "On the minimum audible angle," *J. Acoust. Soc. Am.* **30**, 237–246.
- Mills, A. W. (1960). "Lateralization of high-frequency tones," *J. Acoust. Soc. Am.* **32**, 132–134.
- Ochi, A., Yamasoba, T., and Furukawa, S. (2014). "Factors that account for inter-individual variability of lateralization performance revealed by correlations of performance among multiple psychoacoustical tasks," *Front. Neurosci.* **8**(27), 1–10.
- Remme, M. W. H., Donato, R., Mikiel-Hunter, J., Ballesterio, J. A., Foster, S., Rinzel, J., and McAlpine, D. (2014). "Subthreshold resonance properties contribute to the efficient coding of auditory spatial cues," *Proc. Natl. Acad. Sci.* **111**(22), E2339–E2348.
- Saberi, K., Antonio, J. V., and Petrosyan, A. (2004). "A population study of the precedence effect," *Hear. Res.* **191**, 1–13.
- Shinn-Cunningham, B. G., Zurek, P. M., Durlach, N. I., and Clifton, R. K. (1995). "Cross-frequency interactions in the precedence effect," *J. Acoust. Soc. Am.* **98**, 164–171.
- 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 Haftner, E. R. (2009). "A recency effect in sound localization?," *J. Acoust. Soc. Am.* **125**, 2914–2924.
- Strouse, A., Ashmead, D. H., Ohde, R. N., and Grantham, D. W. (1998). "Temporal processing of the aging auditory system," *J. Acoust. Soc. Am.* **104**, 2385–2399.
- Tobias, J. V., and Schubert, E. R. (1959). "Effective onset duration of auditory stimuli," *J. Acoust. Soc. Am.* **31**, 1595–1605.
- Tollin, D. J. (1998). "Computational model of the lateralization of clicks and their echoes," in *Proceedings of the NATO Advanced Study Institute on Computational Hearing*, edited by S. Greenberg and M. Slaney (IOS Press, Amsterdam), pp. 77–82.
- Wallach, H., Newman, E. B., and Rosenzweig, M. R. (1949). "The precedence effect in sound localization," *Am. J. Psychol.* **62**(3), 315–336.
- Whitmer, W. M. (2004). "The Franssen effect," Ph.D. dissertation, Loyola University Chicago, Chicago, IL.
- Yost, W. A., Mapes-Riordan, D., and Guzman, S. J. (1997). "The relationship between localization and the Franssen effect," *J. Acoust. Soc. Am.* **101**, 2994–2997.
- Yost, W. A., and Soderquist, D. R. (1984). "The precedence effect: Revisited," *J. Acoust. Soc. Am.* **76**, 1377–1383.
- Zurek, P. M. (1980). "The precedence effect and its possible role in the avoidance of interaural ambiguities," *J. Acoust. Soc. Am.* **67**, 952–964.
- Zurek, P. M. (1987). "The precedence effect," in *Directional Hearing*, edited by W. A. Yost and G. Gourevitch (Springer-Verlag, New York), pp. 85–105.
- Zwislocki, J., and Feldman, R. S. (1956). "Just noticeable differences in dichotic phase," *J. Acoust. Soc. Am.* **28**, 860–864.