

A recency effect in sound localization?

G. Christopher Stecker^{a)}

Department of Speech and Hearing Sciences, University of Washington, 1417 NE 42nd Street, Seattle, Washington 98105

Ervin R. Hafter

Department of Psychology, University of California at Berkeley, 3210 Tolman Hall No. 1650, Berkeley, California 94720

(Received 10 June 2008; revised 17 March 2009; accepted 31 March 2009)

In a free-field pointing task, listeners localized trains of 4–32 spatially distributed Gabor clicks (narrowband impulses) centered at 4-kHz carrier frequency and repeating at an interval of 5 ms. Multiple regression coefficients estimated the perceptual “weight” applied to each click in a train during location judgments. Temporal weighting functions obtained in this way exhibited two key features: onset dominance, as evidenced by high weight on the initial click, and “upweighting” of late-arriving sound, as evidenced by weights that gradually increased over the duration of each click-train. Across all tested click-train durations, and despite randomly varying the durations from trial to trial, the greatest post-onset weights were consistently found for clicks at or near the offset. The results imply a special importance of late-arriving sound rather than feedforward recovery from onset dominance, and are broadly consistent with recency effects resulting from temporal integration. © 2009 Acoustical Society of America. [DOI: 10.1121/1.3124776]

PACS number(s): 43.66.Qp, 43.66.Pn, 43.66.Mk [RLF]

Pages: 3914–3924

I. INTRODUCTION

According to Rayleigh’s duplex theory (Strutt, 1907), azimuthal localization of high-frequency ($> \approx 2$ -kHz) sound relies primarily on interaural level differences (ILDs); sensitivity to interaural time difference (ITD) is reduced at these frequencies, due in part to ambiguous correspondences of tone phase at the two ears. For modulated high-frequency sounds, however, localization may utilize ITD carried by temporal envelopes (“envelope ITD”). Over the past several decades, numerous studies have investigated important questions such as whether sensitivity to envelope ITD is similar or inferior to fine-structure ITD at low frequencies and over what range of modulation frequencies envelope-ITD sensitivity is evidenced (see, e.g., Henning, 1974; McFadden and Pasanen, 1975, 1976; Nuetzel and Hafter, 1976). Two relatively recent findings are of special interest. First, envelope-ITD sensitivity at high frequency can be similar to fine-structure ITD at low frequency, at least for sounds with temporally interrupted modulators, such as trains of filtered impulses (Hafter and Dye, 1983; Hafter and Buell, 1990) or “transposed tones” (van de Par and Kohlrausch, 1997). Second, psychophysical and physiological sensitivity to ongoing envelope ITD in high-frequency sounds is significantly impaired at modulation rates above 150–200 Hz. At these rates, discrimination of ongoing envelope-ITD information becomes difficult or impossible in the absence of onset ITD (e.g., when the stimulus is slowly gated on at both ears; Bernstein and Trahiotis, 2002). When onset cues are available, this rate-limitation manifests as increased reliance on

onset ITD and diminished influence of ongoing ITD (Hafter and Dye, 1983; Saberi, 1996). The affected modulation rates and other characteristics of this *onset dominance* (cf. Freyman *et al.*, 1997) call to mind a range of closely related binaural phenomena including binaural adaptation (Hafter and Dye, 1983; Hafter, 1997), the precedence effect (Wallach *et al.*, 1949), and the Franssen effect (Franssen, 1962; Hartmann and Rakerd, 1989).

The relative contributions of onset and ongoing envelope ITD have been assessed using a variety of methods (Tobias and Schubert, 1959; Abel and Kunov, 1983; Buell *et al.*, 1991; Saberi and Perrott, 1995; Buell *et al.*, 2008). A particularly illuminating approach has been to measure the relative sensitivity of listeners’ localization responses to alterations of the spatial cues contained in each temporal portion of the stimulus (e.g., each click in a train of clicks). Plotting the relative sensitivity over time gives a temporal weighting function (TWF) for each stimulus. For example, Stecker and Hafter (2002) presented trains of filtered clicks from loudspeakers in the free-field, varying the location randomly from click to click within each presentation. Following previous studies that used similar techniques to estimate spectral and temporal weights in tone detection and discrimination tasks (Berg, 1989; Richards and Zhu, 1994; Lutfi, 1995; Sadralodabai and Sorkin, 1999), the authors used multiple linear regression of listeners’ pointing responses onto the individual click locations to estimate TWFs for click-trains varying in rate. Consistent with previous and subsequent reports, the TWFs indicated strong onset dominance (in the form of increased weight on click 1) for click rates at or above 200 Hz (interclick intervals [ICIs] ≤ 5 ms). This finding was similar for trains of 2 or 16 clicks and agreed closely with TWFs measured in previous studies of stimuli varying only in ITD (Shinn-Cunningham *et al.*, 1993, 1995;

^{a)}Author to whom correspondence should be addressed. Electronic mail: cstecker@u.washington.edu

Saberi, 1996; Dizon *et al.*, 1998; Stellmack *et al.*, 1999). Different from those prior studies was the additional finding that late portions of the stimulus (clicks at or near the offset) were weighted more strongly than intermediate portions—although a similar result was recently reported for virtual-space localization in the vertical plane (Macpherson and Wagner, 2008). This upweighting of late-arriving sound is intriguing because it suggests that the temporal integration of auditory spatial information could be sensitive to the nature of that information, the behavioral requirements of the experimental task, or both.

Among several candidate explanations for the upweighting of late-arriving sound, the most straightforward is that upweighting simply reflects recovery from onset dominance. In the current report, we describe an experiment that tests this explanation via two closely related hypotheses: a “fixed-recovery” hypothesis in which the recovery time is fixed to a specific duration or number of clicks and an “anticipatory-recovery” hypothesis in which the recovery time adjusts according to the stimulus context. That might allow the system to anticipate stimulus durations when predictable and recover sensitivity to spatial cues contained in the stimulus offset. Note that “anticipation” of stimulus durations might require durations to be highly predictable (e.g., as when the duration is fixed across blocks of trials) or might instead adjust quickly in response to local trial-by-trial variation in stimulus duration. In the first case (global predictability) anticipation should be thwarted (and upweighting reduced) by sufficient randomization of stimulus durations (condition 1 of the current study), while in the second case (local predictability), recovery would not be completely thwarted by randomization but would vary depending on whether stimulus durations do or do not repeat from trial to trial (condition 2). The current study tests these possibilities by measuring TWFs over a range of stimulus durations that vary from trial to trial.

II. METHODS

Details of the experimental apparatus, configuration, and analysis used here were previously described by Stecker and Hafter (2002). Additional details regarding the apparatus and basic temporal-weighting analysis can be found in that paper.

A. Subjects

Four subjects participated in this experiment. One (CS) was the first author; others were paid subjects naive to the purpose of the experiments. Three of the subjects were tested in condition 1 and two in condition 2 (see Secs. II D and II H, below). All subjects had normal audiograms (pure-tone thresholds <10-dB hearing loss (HL)) between 125 and 8000 Hz.

B. Stimuli

Stimuli were trains of narrowband Gabor clicks (Gaussian-windowed tone bursts) sampled at 50 kHz. Each click consisted of a 4-kHz cosine multiplied by a Gaussian temporal envelope with σ equal to 212 μ s.¹ The resulting spectral envelope was Gaussian with $\sigma=750$ Hz (half-

maximal bandwidth ≈ 1.8 kHz). Trains of 4, 8, 16, or 32 such clicks were synthesized with a peak-to-peak ICI of 5 ms. This ICI was chosen for its combination of clear upweighting and similarly clear (though moderate) onset dominance in the data of Stecker and Hafter (2002). Stimulus level, measured from the listener’s position with a continuous train of clicks at 5-ms ICI, was approximately 35-dB sound pressure level (SPL) (A-weighted, “fast” setting) and clearly audible for all subjects.

C. Apparatus

Listeners were seated in an anechoic chamber (Eckel Corp., $8.3 \times 5.4 \times 4.0$ m³), facing an array of 12 ear-height loudspeakers (Audax model MHD12P25 FSM-SQ) separated by 5.5° each and spanning 60.5° in listener-centered azimuth. Loudspeakers were equalized through digital inverse-filtering (frequency distortion within ± 1 dB over the range 2–6 kHz), delay, and attenuation to eliminate spectral and path-length differences between speakers. Loudspeakers were obscured visually by an acoustically-transparent white curtain 42 in. in front of the listener.

D. Stimulus presentation and listeners’ task

On each trial, a location within the loudspeaker array was selected at random. This location defined the center of a group of five adjacent loudspeakers spanning 22° azimuth. (Thus, there were eight unique loudspeaker groups within the 12-loudspeaker array; group centers spanned the central 37.5° of azimuth.) Each of the 4–32 clicks comprising the click-train stimulus was presented from one loudspeaker in the group, selected randomly for each click. This random variation in location across clicks in the train allows the computation of observer weights for each click.

At the beginning of each trial, subjects faced the loudspeaker array with head still as a stimulus was presented from one of the eight potential loudspeaker groups. Next, subjects directed their gaze (without moving their heads) to foveate the perceived location. Next, they recorded a localization response by marking the point of gaze with a hand-directed laser pointer, the orientation of which was recorded digitally and converted to head-centered coordinates for analysis. This procedure was employed in order to maintain a head-centered coordinate system for responses, avoiding confusion related to mapping from laser-pointer orientation to sound-source position in three-dimensional space. In the event that more than one acoustic image was perceived, subjects were instructed to respond to the leftmost image.² Following each response, subjects returned the laser spot to “home” position (head-centered elevation $>28^\circ$, azimuth $0^\circ \pm 17^\circ$) to initiate the next trial after a 1-s delay. Each experimental run consisted of 100 uninterrupted trials. Subjects were allowed to take breaks between runs and completed between 7 (subject LS) and 16 (all other subjects) runs per condition.

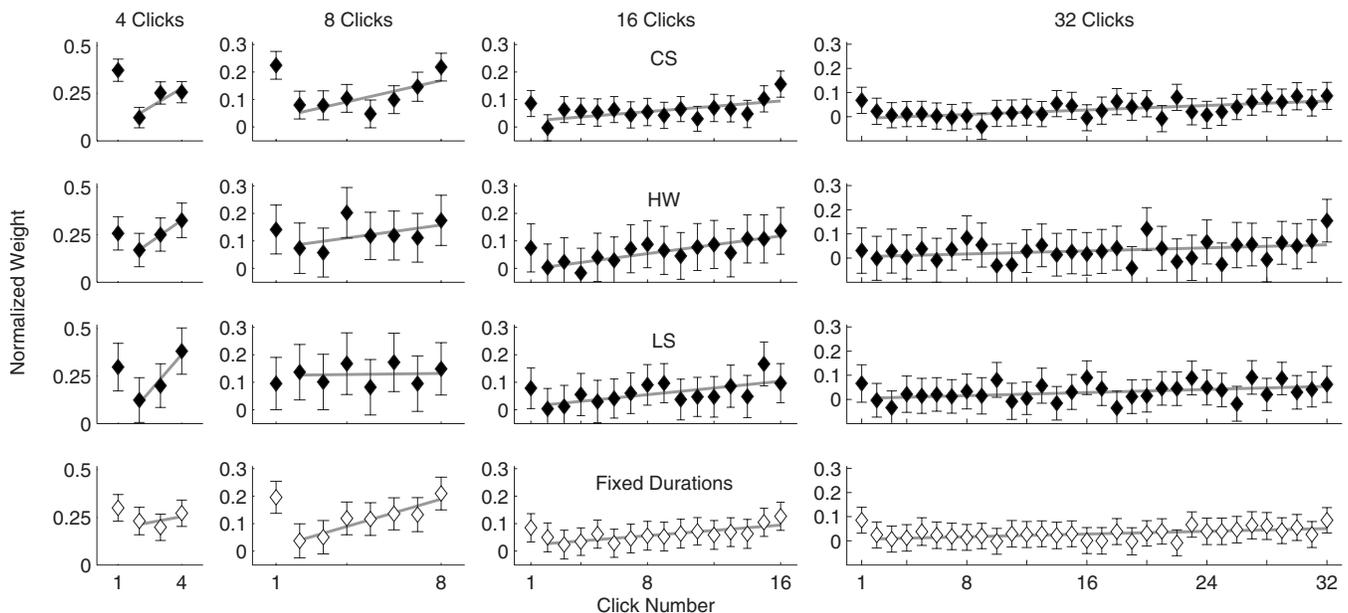


FIG. 1. Normalized weights for click-trains. Upper three rows (filled symbols) plot TWFs for individual subjects, with train length N varied from trial to trial (condition 1). Bottom row (open symbols) plots comparison data obtained with train length N fixed across trials in a block (adapted from Stecker, 2000; Stecker and Hafter, 2002). Columns plot TWFs for trains of 4 (left column), 8, 16, or 32 (right column) clicks. Error bars indicate 95% confidence intervals on weight estimates, and gray lines indicate best linear fit to weights of clicks 2– N .

E. Temporal weighting analysis

The perceptual weights applied to each click in a train were estimated using multiple linear regression of listener response azimuth θ_R onto the azimuths of individual clicks θ_i :

$$\hat{\theta}_R = \sum_{i=1}^N b_i \theta_i + k. \quad (1)$$

For comparison across subjects and conditions, regression coefficients b_i were normalized to sum to 1 over each stimulus duration

$$w_i = \frac{b_i}{\sum_{j=1}^N b_j} \quad (2)$$

with the resulting normalized weights w_i indicating each click's *relative* influence on the listener's response. Individual weights thus vary from 0 (indicating no linear relationship between click location and response) to 1 (indicating a strong linear relationship). Plots of w_i weights vs click number comprise the TWFs and indicate how click effectiveness varies over the stimulus duration. Figures 1 and 3 plot normalized TWFs for each listener, along with 95% confidence intervals on each normalized weight (Stecker and Hafter, 2002). Figure 5 plots across-subject means of these normalized TWFs.

A reasonable strategy for this task would place equal weight on each click since all clicks are equally informative with regard to the stimulus location. For a discrimination task, such a strategy can be considered *optimal* because the effective variance of a location estimate is reduced by equal averaging across multiple clicks in a train (Saberi, 1996). Unequal weighting, in contrast, is sub-optimal because it reduces this benefit (effectively averaging over a smaller num-

ber of clicks). A similar argument for the effectiveness of an equal-weighting strategy can be made for localization in the free-field, although “optimality” is not well-defined for this task.

F. Measures of upweighting

We quantified the degree of upweighting in each TWF in two different ways. The first approach was based on the *average ratio* (AR) originally defined by Saberi (1996) as the ratio of onset click weight to the average of post-onset click weights. We chose to use AR over Saberi's *precedence ratio* (ratio of onset weight to sum of remaining weights) because AR more appropriately facilitates comparison across stimuli differing in click number. For the purposes of this study, we redefined AR as the ratio of onset or offset weight to the mean of intermediate weights (i.e., the mean excluding onset and offset clicks):

$$\text{AR}_{\text{onset}} = \frac{w_1}{\sum_{i=2}^{N-1} w_i / (N-2)} \quad (3)$$

or

$$\text{AR}_{\text{offset}} = \frac{w_N}{\sum_{i=2}^{N-1} w_i / (N-2)}. \quad (4)$$

The second approach we used to quantify upweighting reflects the apparently gradual increase in weights reported by Stecker and Hafter (2002). We used linear regression to fit straight lines through weights 2– N , separately for each subject and condition. Residuals were examined for approximately uniform distribution and did not indicate any clear nonlinear trends. For plotting, slope values were converted to *rise* (slope multiplied by the range of clicks used for fitting), an estimate of the difference between offset and click 2 that

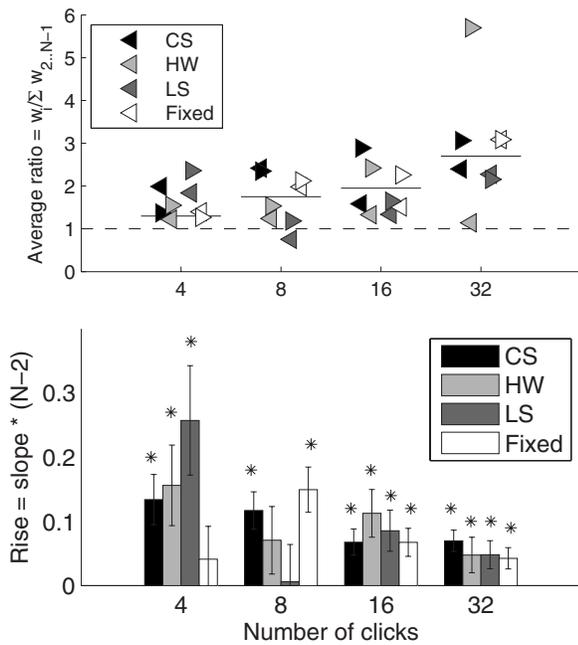


FIG. 2. Top: AR, quantifying the relative influence of onset (left-pointing triangles) and offset (right-pointing triangles) clicks, is plotted for each combination of train-length N (x -axis) and subject tested in condition 1 (filled symbols). Comparison data from fixed-duration studies (Stecker, 2000; Stecker and Hafter, 2002) are plotted with open symbols. The dashed line represents a ratio of 1 (the value that would obtain for equal sensitivity across all clicks in a train); symbols falling above the short solid lines indicate AR significantly greater than 1 at the $p < 0.05$ level based on parametric bootstrap testing. Bottom: rise, defined as the expected weight change (based on a linear fit) from click 2 to click N (offset), is plotted for each combination of N and subject tested in condition 1 (filled bars) and for fixed-duration studies (open bars). Error bars indicate standard error of the rise estimate, while asterisks (*) indicate values significantly greater than 0 ($p < 0.05$).

would occur given a perfectly linear trend (hence, rise underestimates nonlinear upweighting confined to clicks near the stimulus offset).

G. Null-hypothesis testing

Statistical evaluation of AR and rise employed a parametric bootstrap procedure to evaluate the null hypothesis that $AR=1$ and $rise=0$, separately for each combination of subject and condition. Bootstrap weights were generated for each click in a train by random sampling from a normal distribution defined by the original estimate and standard error for that click weight. AR and rise were computed for each of 10 000 bootstrapped TWFs comprised of these weights to generate sampling distributions for each statistic. Sampling distributions for AR were approximately log normal, and the significance level of each AR estimate was computed empirically from the proportion of bootstrapped $AR \leq 1$. Sampling distributions for rise were normal. Standard error (e.g., error-bars in Fig. 2) was computed directly from the distribution for each rise estimate and significance levels computed from the proportion of bootstrapped $rise \leq 0$. Results of significance testing are displayed in Figs. 2 and 4 for each combination of subject and condition.

A second level of significance testing compared values of AR and rise across subjects and conditions. For this, we used Wilcoxon’s one-sample signed-rank test (MATLAB statistics toolbox function “signrank”) to evaluate the null hypotheses that (1) median AR equals 1 across subjects and conditions and (2) median rise equals 0. We used the paired-samples version of the same test to compare AR_{onset} and AR_{offset} across subjects and conditions.

H. Train-length manipulation

The length of each click-train was selected randomly on each trial. In condition 1, train length was selected randomly to be 4, 8, 16, or 32 clicks in order to test the hypotheses that (1) weight increases occur after a fixed number of clicks (the “fixed recovery” hypothesis) or (2) weight increases occur at different latency based on “anticipation” of a predictable train offset (the “global anticipatory recovery” hypothesis). In condition 2, train length was constrained to be either 8 or 16 clicks to increase the likelihood of train-length repeats on successive trials so that weighting functions could be reliably computed for both repeated and non-repeated train lengths. This condition tests a “local” anticipatory-recovery hypothesis, in which weight increase occurs at a time matching the duration of the immediately previous trial. Subjects CS, HW, and LS were tested in condition 1, and subjects HW and TL were tested in condition 2.

III. RESULTS

A. Localization accuracy and reliability

The accuracy and reliability of listeners’ localization responses were quantified using the multiple regression model responsible for temporal weight estimation [Eq. (1)]. For the current study, this approach offers two key advantages over a more traditional approach in measuring listener performance. First, due to the randomization of individual click locations and the unequal temporal weighting applied to click-trains, it was not possible to compute a “correct” location for each trial outside the context of the regression model. Second, the reliability of localization responses could be computed from the proportion of variance explained by the model (R^2), thus taking into account the effects of individual and condition-specific differences in localization gain, bias, and temporal weighting.

Reliability. R^2 ranged from 0.81 to 0.96, with a mean of 0.89 and standard deviation of 0.04, across subjects and conditions, indicating good reliability of localization responses for all subjects. Note that the quantity $1-R^2$ (proportion of localization variance not explained by the model) estimates the localization error variance³ as the deviation from expected response given the full model, thus taking account of intersubject and interstimulus differences in localization accuracy.

Accuracy. Because the main focus of the study was on the relative effects of individual click locations on listeners’ responses (i.e., TWFs are not affected by gain or bias of listener’s responses) and because the task itself was subjective (in the sense that listeners pointed to perceived locations without feedback), absolute accuracy measures are not par-

ticularly meaningful but are included here. Localization gain (the ratio of expected response azimuth to stimulus azimuth) was estimated by the sum of (non-normalized) regression weights b_i over all clicks. Gain ranged from 0.66 to 1.23, with a mean of 0.98 and standard deviation of 0.15, across subjects and conditions. Bias, estimated by the constant term k of the regression equation, ranged from $+0.7$ to $+8.5$ (positive value indicates rightward bias), with a median of $+3.25^\circ$ azimuth across subjects and conditions (subject HW exhibited a consistent rightward bias of 6° – 8.5° in all conditions; other subjects' biases were generally close to $+3^\circ$). Gain near 1 and bias near 0 are indicative of accurate (though not necessarily reliable) localization.

B. Temporal weighting functions

Figure 1 plots the TWFs obtained in condition 1 (filled symbols plot normalized weights for each click, with error bars indicating 95% confidence intervals). Although some individual variability is present, key features of the TWFs—specifically, the greater weight for clicks at onset and near offset than for intermediate clicks—appear for most subjects and train lengths. For example, in all but two cases (subject LS at 8 and 32 clicks), the weight applied to click 2 fell below the 95% confidence interval on weight for click N . In all but one case (LS, eight clicks), weights can also be seen to gradually increase over the post-onset duration of each click-train, as reported by Stecker and Hafter (2002) for 16-click-trains. TWFs from that study, along with data for 4-, 8-, and 32-click-trains (adapted from Stecker, 2000) are plotted for comparison in the bottom row of Fig. 1 (open symbols). Those studies measured TWFs for click-train durations that remained fixed across blocks of trials and used a slightly different loudspeaker arrangement, but were otherwise similar to the current study and tested five subjects that included those of the current study. Note that only data for 5-ms ICI are included in the figure. Again, key features of the TWFs (onset dominance and upweighting) appear similar across train lengths and regardless of train-length randomization.

Values of the AR for each subject and condition are plotted in Fig. 2 (top panel). AR values varied somewhat across subjects and train length N , but were consistently and significantly greater than 1 for both onset (Wilcoxon one-sample signed-rank test, $T=4$, $p<0.005$) and offset ($T=0$, $p<0.0005$) clicks, indicating greater influence of these clicks than intermediate clicks. Furthermore, AR of the offset click was slightly but significantly greater than that of the onset click ($T=11$, $p<0.05$). That difference suggests that the magnitude of upweighting is not limited to the magnitude of onset dominance, as would be expected if upweighting were simply recovery from onset dominance.⁴

The lower panel in Fig. 2 plots the rise in weights from click 2 to click N for each combination of train duration and subject, as measured by the slope of best-fitting linear function. Overall, the rise was significantly positive, ($T=0$, $p<0.0005$), similar between subjects, and declined slightly with increasing train length N . The two measures of upweighting (AR and rise) are complementary in that AR is especially sensitive to sharp weight increases near the end of

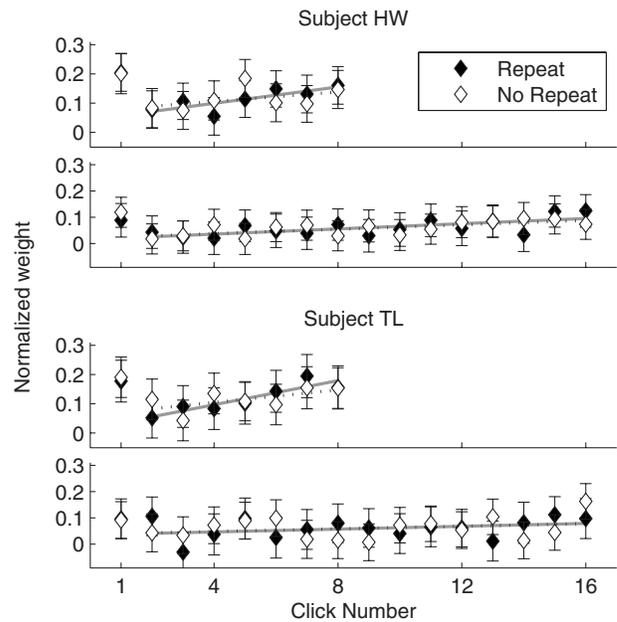


FIG. 3. Normalized weights for click-trains with train-length varied from trial to trial (condition 2). For each subject, upper and lower panels plot functions for 8- and 16-click-trains, respectively. Filled symbols represent weights from trials which repeated the train-length of the immediately-previous trial; open symbols represent weights from trials where train-length differed from the previous trial. Error bars indicate 95% confidence intervals on weight estimates. Lines indicate best linear fit to click weights $2N$ for repeated (gray lines) and non-repeated (dotted lines) trials. There were no significant differences between repeated-length and non-repeated-length trials.

the sound, whereas rise better characterizes gradual weight increases. Both features are present in the data (note in Fig. 1, for example, the tendency for offset clicks to receive weight in excess of the linear trend), however, and the two measures agree. The measures varied slightly, and in opposite ways, with increasing N , suggesting relatively more abrupt weight increase at offset (vs gradual increase) for longer trains. Overall, weights were consistently greater for clicks presented during later portions than middle portions of the train.

Figure 3 plots TWFs obtained in condition 2; consistent with the findings of condition 1, these exhibited both onset dominance and upweighting of later clicks. TWFs for repeated-length trials (filled symbols) were not significantly different from TWFs for non-repeated trials (open symbols), indicating little if any effect of the local stimulus context on the form of temporal weighting. That result is corroborated by AR values exceeding 1 and rise values exceeding 0 in Fig. 4. Values were similar overall to those measured in condition 1 for both repeated and non-repeated trials. Estimates of rise did not differ significantly between repeated and non-repeated train lengths based on permutation bootstrap testing (8 clicks: $p<0.11$, 16 clicks: $p<0.42$).

IV. DISCUSSION

A. Prior studies indicating recovery from onset dominance

The upweighting effect described in this paper demonstrates one way in which spatial cues contained in the later,

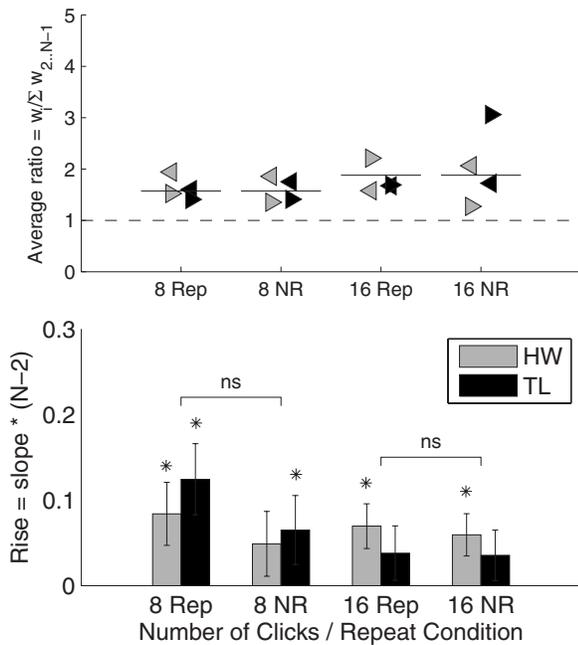


FIG. 4. AR and rise values for data obtained in condition 2. Format identical to Fig. 2, but with AR and rise plotted separately for 8- and 16-click-trains immediately preceded by identical-length trains (Rep=“repeat”) or by different-length trains (NR=“no repeat”). Comparison lines (“ns”) indicate lack of significant difference ($p > 0.05$) between rise measured with repeated or non-repeated-length trains.

ongoing, portions of a stimulus contribute to sound localization judgments. The role of such ongoing cues has been investigated in several past studies. Tobias and Schubert (1959), for example, studied the duration of ongoing sound necessary to counteract transient cues present at sound onset and offset. Based on their results, Tobias and Schubert (1959) argued that ongoing cues completely dominate the localization of broadband sounds longer than 100 ms in duration; for shorter stimuli, transient and ongoing cues trade. Because the transient cues employed in that study included both onset and offset cues in agreement with each other, the results actually demonstrated a dominance of intermediate over onset and offset cues—contrary to the current results—with long-duration stimuli. The consequent prediction that both onset dominance and upweighting should be eliminated for stimuli over 100 ms in duration is not strongly supported by the current results⁵ (Fig. 2), although casual inspection of TWFs plotted in Fig. 1 might suggest somewhat flatter functions for longer stimuli.

Freyman *et al.* (1997) investigated the apparent discrepancy between results like those of Tobias and Schubert (1959) and other studies that demonstrated strong onset dominance (e.g., Kunov and Abel, 1981; Saberi and Perrott, 1995). Based on their results, Freyman *et al.* (1997) argued that onset dominance arises in cases where stimulus periodicity or spectral sparseness renders ongoing cues ambiguous. Note that the high-rate Gabor click-trains used in the current study contain both types of “ambiguous” features and demonstrate strong onset dominance, consistent with the results of Freyman *et al.* (1997). This likely explains the discrepancy between current results and the ongoing-cue dominance reported by Tobias and Schubert (1959).

Zurek (1980) asked listeners to discriminate the lateral positions of pairs of noise bursts and found binaural discrimination to be impaired during a period 2–5 ms following the initial stimulus onset. He suggested that sensitivity to binaural information is transiently reduced following onset and recovers over the subsequent 10 ms. In one experiment, Zurek (1980) inserted a brief (5-ms) dichotic noise “probe” into a longer (50-ms) diotic noise burst (thus introducing a leading and a trailing diotic “fringe”). Discrimination performance was strongly affected by the temporal position of the probe within the background: best near the background onset or offset and worst approximately 2–5-ms post-onset. Discrimination performance improved monotonically with delay beyond this value, reaching near-normal levels only close to the offset of the background noise, somewhat similar to the TWFs measured in the current study. Akeroyd and Bernstein (2001) extended this finding by repeating Zurek’s (1980) experiment with added conditions that omitted either the leading or trailing fringe. Consistent with the results of Zurek (1980) and those of the current experiment, they found greatest sensitivity (lowest thresholds) when the probe occurred near the onset or offset of the background noise burst (e.g., when either the leading or trailing burst was omitted).

B. Evaluation of recovery hypotheses

The results of Zurek (1980) provide one candidate explanation for the upweighting effect considered in the current study: that upweighting reflects recovery from the temporary effects of onset dominance. Here, we consider that explanation with respect to results of the current study. Specifically, we consider two sets of hypotheses: (1) the fixed recovery hypothesis, in which recovery takes place over a fixed time course and (2) the anticipatory recovery hypothesis, in which recovery takes place over a time course that is adjusted in a predictive fashion. Two types of anticipatory mechanisms are considered. The “global” anticipatory hypothesis posits that recovery time is adjusted to match the expected stimulus duration when durations are highly predictable (e.g., when they do not change from trial to trial), while the “local” anticipatory hypothesis posits more rapid adjustment of recovery time on a trial-by-trial basis.

Following Zurek (1980), the fixed recovery hypothesis assumes that the post-onset reduction in weights (i.e., onset dominance) reflects a temporary insensitivity to binaural information and that the binaural system recovers sensitivity given sufficient time. The recovery time might reflect the time constant of post-onset inhibition (Zurek, 1980; Lindemann, 1986) or to increasing likelihood of “restarting” the binaural adaptation mechanism (Hafer and Buell, 1990). Previously, Stecker and Hafter (2002) reported recovery of weights to occur near the offsets of 16-click-trains across a range of ICI from 3 to 8 ms and thus a range of train durations spanning 47–122 ms (Stecker and Hafter, 2002). That result supports the view that recovery time should be expressed in number of clicks (16) rather than elapsed time.⁶ The fixed-recovery hypothesis predicts recovery to take place at the same point in each train, regardless of train length. Thus, recovery is expected to occur near the midpoint

of 32-click-trains, but remain incomplete for trains of fewer than 16 clicks. The TWFs plotted in Fig. 1 are not consistent with this prediction, as recovery appears consistently near the offset of each tested click-train, regardless of length. A corollary prediction of the fixed-recovery hypothesis is that upweighting should be significantly reduced or eliminated for shorter-duration trains, a prediction that is not supported by either of the summary statistics plotted in Fig. 2.

As an alternative to the fixed-recovery hypothesis, we next consider the anticipatory recovery hypothesis. By this account, upweighting reflects recovery from onset dominance with an adjustable—rather than fixed—recovery time. For example, since echoes act to prolong the proximal stimulus in a room-specific manner, such adjustment could be a means to optimize the temporal extent of onset dominance (echo suppression). Indeed, this trial-to-trial adaptation would be reminiscent of the “buildup” of precedence effects with repeated stimulation (Clifton and Freyman, 1989). The anticipatory-recovery hypothesis predicts recovery to occur at or near the end of each click-train if the duration is predictable from the context of recent stimuli. TWFs measured with duration fixed across blocks of trials (Stecker, 2000; Stecker and Hafter, 2002, lower panels of Fig. 1) appear consistent with that prediction. In the current study, we reduced predictability by randomizing train durations in condition 1 but noted a similar pattern of recovery near the end of each stimulus regardless of duration. Thus, upweighting does not appear to require predictable train lengths, contrary to the prediction of the anticipatory-recovery hypothesis.

While the results obtained in condition 1 argue against an anticipatory recovery mechanism requiring blocks of fixed train durations (a global anticipatory hypothesis), a remaining possibility is for the recovery time to be adjusted on a more rapid, trial-by-trial, basis (a local anticipatory hypothesis). In that case, stimulus durations are predictable when they match the duration of the immediately previous trial (“repeated trials”). TWFs for repeated trials are thus predicted to show recovery near the stimulus offset (as for fixed durations), while TWFs for non-repeated trials should experience recovery at some other time (e.g., reduced recovery if the previous trial was longer in duration, or premature recovery if the previous trial was shorter). In condition 1, repeated trials accounted for 25% of trials, so one might expect some effect of local predictability on TWFs. Casual examination of Fig. 1 suggests that some TWFs exhibited slight weight increases near clicks 4, 8, and 16 in trains of other lengths, possibly consistent with the local hypothesis. Condition 2 separately tested this possibility, however, and found no significant differences between TWFs measured with repeated and non-repeated trials. Thus, despite the suggestive features of condition-1 TWFs, the results are not consistent with either local or global anticipatory recovery.

Overall, TWFs measured in this study showed increased weights near the end of each click-train stimulus, consistent with the observations of Stecker and Hafter (2002). This pattern was observed consistently regardless of click-train duration and despite randomization to reduce the predictability of stimulus duration. These results therefore strongly suggest that upweighting does not reflect a feedforward recovery

from onset dominance with either a fixed or anticipatory recovery time. Rather, the weight increases reflect a specific contribution of late-arriving sound, possibly independent of onset dominance. In other words, sound localization reflects a *post hoc* combination of multiple cues (Sabetri, 1996) rather than a strictly feedforward accumulation of spatial information.

C. Prior studies that failed to show upweighting

In addition to the current study, a number of other recent studies have also measured TWFs for auditory spatial cues. Although these studies have generally been in agreement with respect to evidence for onset dominance, upweighting has not been consistently reported. Notably, two studies (Sabetri, 1996; Dizon *et al.*, 1998) measured TWFs for ITD using headphone presentation, finding strong evidence for onset dominance at fast rates, but no suggestion of upweighting. In contrast, the current study, Stecker and Hafter (2002), and a recent study of TWFs for stimuli varying in median-plane elevation (Macpherson and Wagner, 2008) all reported upweighting of late-arriving sound. What similarities and differences among these studies might account for the differences in findings?

Difference 1: The types of spatial cues presented. Among the most obvious differences are the spatial cues available to listeners in each study. Stecker and Hafter (2002) presented stimuli in the free-field and Macpherson and Wagner (2008) used virtual acoustic space (VAS) recordings based on the acoustic effects of listeners’ own ears in the free-field; the studies of Sabetri (1996) and Dizon *et al.* (1998) employed headphone presentation, with spatial cues limited to ITD only.⁷ Free-field and VAS stimuli possess cues related to ILDs and direction-dependent spectral cues generated by the acoustic effects of the head and outer ears, along with corresponding ITD cues. Varying only ITD, in contrast, places these cues in conflict with one another (ILD and spectral cues indicate a constant, near-midline azimuth). The relative weighting of ITD and ILD cues has been shown to vary with the listening situation (Rakerd and Hartmann, 1985), and ambiguity (as in a cue-conflict situation) may cause listeners to de-emphasize one or the other cue. Two possibilities are suggested by the differences in spatial-cue content. One possible explanation is that upweighting reflects mainly sensitivity to late-arriving ILD and/or spectral cues and somehow does not apply to the temporal integration of ITD. A second possibility is that conflict between cues reduces upweighting by rendering late-arriving cues ambiguous or uninformative (Rakerd and Hartmann, 1985; Freyman *et al.*, 1997).⁸

Difference 2: The nature of tasks involved. Another important difference among these studies is the type of response employed. Dizon *et al.* (1998), for example, used an adjustment technique with acoustic pointer to indicate the perceived lateral positions of stimuli. In that study, test and pointer alternated until subjects were satisfied with the match (a closed-loop procedure in which listeners have multiple opportunities to compare stimuli and minimize adjustment error). Sabetri (1996) had subjects discriminate the lateral po-

sitions of stimuli as left or right of midline. In the current study (as in [Stecker and Hafter, 2002](#); [Stecker, 2000](#)) subjects made a saccadic eye movement following presentation of a single stimulus (an open-loop procedure with no opportunity for “fine-tuning” the response over multiple presentations). [Macpherson and Wagner \(2008\)](#) had subjects turn their heads in the direction of the perceived location, also in open-loop fashion. The tasks employed in the latter two studies are the most similar; both used orientation (pointing to a location in space) in an open-loop procedure. The discrimination procedure used by [Saber \(1996\)](#) is most clearly different; discrimination does not necessarily require that subjects be able to report a perceived location, and engages a motor response (button press) that does not require precise positioning within the spatial reference frame of the stimulus representation. The case of [Dizon et al. \(1998\)](#) is intermediate; subjects were asked to precisely judge the spatial location of the stimulus, as in the current study, but they did so in the context of a closed-loop task. The various procedures have different memory demands that could explain the difference in upweighting (see Sec. IV D 1). Alternately, we might consider the differences in neural mechanisms responsible for generating the specific motor responses in each task (see Sec. IV D 3): both of the orienting tasks require a response in the same spatial reference frame as the stimulus location, whereas the adjustment and discrimination tasks require a more arbitrary mapping from perception to motor action.

D. Mechanisms that might contribute to upweighting

Although the results clearly suggest a *post-hoc* analysis of late-arriving sound, the mechanisms underlying that *post-hoc* analysis are not clear. Potentially relevant mechanisms, addressed below, relate to the integration of recent information in sensory memory, the perception of apparent motion, and/or the generation of motor responses.

1. Recency in sensory memory

One possible explanation for the relative importance of late-arriving spatial cues may be found in the nature of memory mechanisms involved in integrating these cues over time. Well known from memory tasks (such as free recall of words from a memorized list), for example, are “recency” effects, whereby memory for recent items is better than for prior items ([Glanzer and Cunitz, 1966](#)). Within the domain of auditory perception, [Sadralodabai and Sorkin \(1999\)](#) noted clear recency effects in listener weighting of temporal intervals in a tone-pattern discrimination task. Although numerous mechanisms have been considered for recency effects at long time scales (seconds to minutes), one view that is especially amenable to shorter time scales is of recency as a general consequence of temporal integration. For example, a leaky integrator combines new information with a decaying representation of prior information, thus acting as a rudimentary memory that emphasizes the recent over the less-recent.

Similarly, [Akeroyd and Bernstein \(2001\)](#) used temporal integration to account for the dependence of ITD and ILD thresholds on varying-length fringes of diotic noise sur-

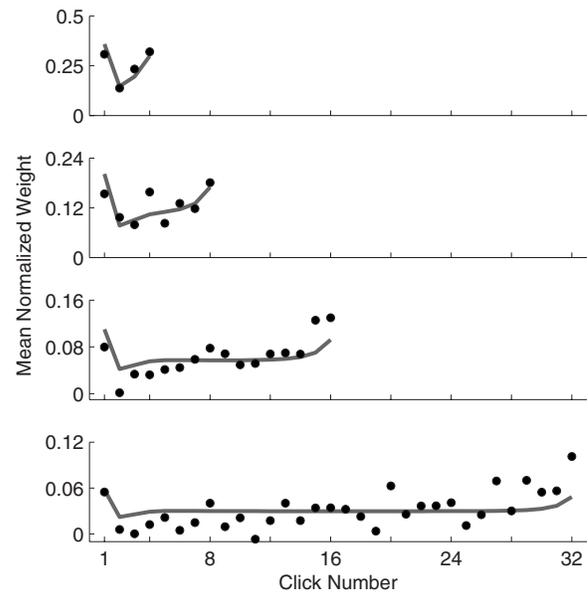


FIG. 5. TWFs predicted by quantitative modeling of binaural temporal integration and post-onset weighting ([Akeroyd and Bernstein, 2001](#)). Circles: mean normalized TWFs across subjects for train lengths tested in condition 1. Solid lines: model predictions (see Appendix for modeling details). The combination of post-onset weighting and binaural temporal integration closely matches the measured TWFs for short stimuli (four to eight clicks). For longer durations, the model accounts for the basic TWF shape but appears to underestimate both the magnitude and the duration of upweighting.

rounding a dichotic noise burst ([Zurek, 1980](#)). Their model combines a temporal integrator (a “binaural temporal window”) with a post-onset weighting function ([Houtgast and Aoki, 1994](#)) to model both onset dominance and upweighting-like effects in their discrimination data. Because the temporal window weights prior inputs by a decaying exponential (as in the leaky integrator), its memory-like behavior similarly emphasizes recent input, resulting in increased weight (and reduced thresholds) for binaural information occurring near the offset of a sound.

In order to quantify the effects of temporal integration on TWFs, we applied the model of [Akeroyd and Bernstein \(2001\)](#) to stimuli as tested in condition 1 of the current study. Details of the model are given in the Appendix. The model predictions are plotted along with cross-subject average TWFs in Fig. 5. The parameters were identical to those used by [Akeroyd and Bernstein \(2001\)](#) to model ITD discrimination thresholds of broadband noise, but still account reasonably well for the narrowband localization-based TWFs of the current study. Both onset dominance and upweighting are apparent in the model predictions. The former is a consequence of the model’s post-onset weighting [Eq. (A1)], while the latter reflects temporal integration [Eqs. (A2) and (A3)]. Intermediate weights are reduced due to “dilution” of binaural information by neighboring clicks; this dilution is reduced at stimulus onset and offset when fewer neighbors fall within the limits of the temporal window. Note that for long durations (16–32 clicks), the model TWFs feature constant weights during the middle of the stimulus and a brief increase near the stimulus offset. The temporal extent of the increase reflects the time constant of the binaural temporal window and in this case appears to underestimate the up-

weighting measured in this study. A longer (i.e., more “sluggish;” [Grantham and Wightman, 1978](#)) temporal window might have provided a better fit to the present data. If so, that result would be consistent with the view that localization involves relatively slower integrative mechanisms than does binaural discrimination.

Note that the possibility of recency effects in sound localization (due to binaural temporal integration or otherwise) does not speak to the issue of why some studies have reported upweighting (this study, [Zurek, 1980](#); [Akeroyd and Bernstein, 2001](#); [Macpherson and Wagner, 2008](#)), while other similar studies have not ([Saber, 1996](#); [Dizon et al., 1998](#)). A satisfactory explanation would require addressing how the use of memory systems differs between these studies. One possibility is that the time constant of temporal integration differs between ILD (or spectral cues) and ITD. Another is that some tasks, such as open-loop orientation, depend on decaying spatial representations that are not utilized in discrimination or closed-loop tasks (see Sec. IV D 3). At this point, however, neither of these possibilities provides a compelling account of the data, insofar as the results of [Zurek \(1980\)](#) and [Akeroyd and Bernstein \(2001\)](#) suggest upweighting for *discrimination* of both ITD and ILD, and the current results demonstrate upweighting for *localization* in the free-field.

2. Apparent motion

Another possibility is that the dynamic locations of stimuli resulted in a perception of apparent motion. While the likelihood of coherent motion of up to 32 clicks is quite low, saltatory motion (e.g., motion caused by displacement of the offset relative to the onset, [Grantham, 1997](#)) could produce a sensation of motion in the direction of the displacement. If listeners’ judgments were, in turn, biased in the direction of motion, we would expect to observe increased weights at or near the offset of the sound, regardless of duration. These saltatory motion cues thus represent another way in which *post-hoc* analysis of spatial information might produce upweighting. However, the argument for upweighting as a result of apparent motion is a bit circular: if the perceived location of a stimulus is dominated by the onset and offset, then motion may be perceived when onset and offset differ in location. Does the motion enhance the effect of the offset click, or vice versa? Furthermore, it does not explain the differences between studies noted above; all of the referenced TWF studies included dynamic locations that could have introduced apparent motion, unless the types of spatial cues strongly determine the salience of saltatory motion (cf. “binaural sluggishness,” [Grantham, 1984](#)). Future experiments that directly manipulate perceived motion may be necessary to evaluate this possibility.

3. Response generation

Finally, the *post-hoc* integration of spatial information suggested by the current results could imply a role for decision and/or response aspects of the task. That is, the observed effects might be strongly response-dependent, reflecting non-auditory (cognitive or motor) processes that are not funda-

mentally involved in the early processing of auditory spatial cues. Could late-arriving cues somehow bias the response representation (e.g., a spatial-motor plan) rather than the sensory representation itself? This view is supported by the apparent failure of upweighting in discrimination and closed-loop adjustment tasks ([Saber, 1996](#); [Dizon et al., 1998](#)), where the two representations are not required to share a common spatial reference frame. Future studies might test this possibility directly by comparing TWFs for identical stimuli across response methods such as open-loop orientation and spatial discrimination.

It is an intriguing possibility that late-arriving cues might affect some tasks and not others, but this explanation requires that response-generation (motor-planning) mechanisms gain access to information about multiple spatial cues (i.e., onset and ongoing cues) *prior to* integration of spatial information across cue type. That is, multiple representations of auditory spatial information (e.g., representations of onset vs ongoing or ITD vs ILD cues) must be separately maintained to relatively high levels within the auditory pathway (that is, at least to those levels involved in generating motor responses). Though difficult to reconcile with a traditional modular view of sensory and motor processing, this interpretation is, in fact, consistent with recent electroencephalographic and lesion-based evidence for separate representations of ITD and ILD within human auditory cortex ([Tardif et al., 2006](#); [Yamada et al., 1996](#); [Schroger, 1996](#); [Ungan et al., 2001](#)). The notion that multiple “channels” of spatial information might be integrated at a very late stage of processing is, furthermore, consistent with current views of how auditory space is represented within auditory cortex. Neurophysiological evidence from animal models suggests that auditory spatial codes are distributed across a small number of broadly tuned sub-populations of cortical neurons ([Stecker and Middlebrooks, 2003](#)) rather than by sharply tuned individual neurons. Although their aggregate spatial information is sufficient to underlie normal behavior, that information might not be integrated across these sub-populations until the level of multimodal and/or sensorimotor integration ([Stecker et al., 2005](#)).

V. SUMMARY AND CONCLUSIONS

The following conclusions can be drawn from the current results.

- (1) TWFs for free-field localization of 4000-Hz Gabor click-trains, presented at 5-ms ICI, are characterized by (1) onset dominance, as evidenced by increased weight on the first click, and (2) upweighting of late-arriving sound, evidenced by a gradual increase in weights following the first click and a correspondingly increased weight on the final click.
- (2) Upweighting affects most strongly those clicks nearest the offset of the train, regardless of stimulus duration, predictability of duration, or similarity to recent stimuli.
- (3) Upweighting of late-arriving sound thus does not reflect a recovery from onset dominance. Among eliminated hypotheses are (1) recovery at a fixed post-onset time, (2) recovery at a variable post-onset time adjusted to match

the (predictable) stimulus duration, and (3) recovery at a variable post-onset time adjusted to the duration of immediately previous stimuli.

ACKNOWLEDGMENTS

The authors thank Miriam Valenzuela and Ephram Cohen for assistance in running this study. Erick Gallun, Bruce Berg, David Wessel, and Frederic Theunissen provided helpful comments during the study design. Erick Gallun, Ewan Macpherson, Andrew Brown, Ian Harrington and Rich Freyman, and three anonymous reviewers provided invaluable comments on earlier versions of the manuscript. Virginia Richards suggested the simplifying term “Gabor click” to describe the Gaussian-filtered impulses used in this and other studies. A portion of this work was previously presented in abstract form [Stecker and Hafter (2002). *J. Acoust. Soc. Am.* 111, 2355] and in the first author’s doctoral dissertation. Work supported by NIH R01 DC00087 (E.R.H.) and R03 DC009482 (G.C.S.).

APPENDIX: MODELING DETAILS

The model of Akeroyd and Bernstein (2001) combines a binaural temporal window of integration with a post-onset weighting function (Houtgast and Aoki, 1994).

Model predictions were computed in the following manner.

- (1) Post-onset weights $x(t)$ were computed as a function of post-onset time t of each click in a train, according to the following post-onset weighting function (Houtgast and Aoki, 1994):

$$x(t) = ae^{-t/T_a} + be^{-t/T_b} + 1. \quad (\text{A1})$$

Parameters for Eq. (A1) were set to the values used by Akeroyd and Bernstein (2001) to model ITD discrimination: $a=3.1$, $T_a=2.8$ ms, $b=-2.1$, and $T_b=5.9$ ms.

- (2) Temporally integrated weights $y(t)$ were computed by calculating the ratio of each $x(t)$ weight to the sum of x weights falling within an asymmetric temporal window $\omega(\tau)$ centered on each click time t (Akeroyd and Bernstein, 2001):

$$y(t) = \frac{x(t)}{\int x(t+\tau)\omega(\tau)d\tau}, \quad (\text{A2})$$

where τ represents time relative to the peak of the binaural temporal window function $\omega(\tau)$:

$$\omega(\tau) = \begin{cases} e^{\tau/T_1}, & \tau < 0 \\ e^{-\tau/T_2}, & \tau \geq 0, \end{cases} \quad (\text{A3})$$

and T_1 and T_2 are time constants that define the response of the temporal window to prior and subsequent clicks, respectively. As for Eq. (A1), parameters for Eq. (A3) were set to the values used by Akeroyd and Bernstein (2001) to model ITD discrimination: $T_1=5.2$ ms and $T_2=7.2$ ms.

Model TWFs were computed by normalizing the temporally integrated weights y so that the set of model weights sums to 1 across all clicks in a train. In Fig. 5, normalized

model TWFs are plotted along with mean normalized TWFs measured in condition 1 of the current study.

¹We have previously referred to this duration as “nominally 2 ms,” based on the total envelope duration calculated at an amplitude resolution of 16 bits (Stecker and Hafter, 2002). Despite the difference in nomenclature, the stimuli used in the current study are identical to those of the previous study.

²This instruction was meant to reduce any effects of listener bias when multiple images were present (for example, by electing to respond to the early-sounding image). Consistent pointing to the leftmost image is expected to flatten TWFs overall when multiple images are perceived rather than to enhance any particular element of the TWF. Subjects were not told to expect multiple images, and in this study, no subjects reported consistent appearances of multiple images. In our previous study (Stecker and Hafter, 2002), the appearance of multiple images was primarily confined to conditions of long ICI (14 ms) and were not reported at 5-ms ICI.

³Because $1-R^2$ also includes variance associated with systematic but non-linear trends in the responses, it estimates an upper bound (rather than an unbiased estimate) on the true localization error.

⁴Note that this result is consistent with the moderate onset dominance observed previously for 5-ms ICI. At shorter ICI, greater onset dominance (and thus greater AR for onset clicks) would be expected.

⁵Note, however, the significant differences between stimuli. Tobias and Schubert (1959) presented broadband noises with amplitude fluctuations reduced by peak limiting, whereas the current study employed narrowband stimuli that were amplitude-modulated at a depth of 100%.

⁶In the current study, ICI was fixed at 5 ms so that duration and click number are perfectly correlated.

⁷Saberi (1996) presented 4000-Hz Gabor click-trains—as used in the current study—with ICI ranging 1.8–12 ms and asked subjects to discriminate the left/right position of trains with ITD varying from click to click. Dizon *et al.* (1998) presented broadband noise bursts divided into four to six temporal “slices” of 2–10 ms duration, each with a randomly selected ITD. Subjects were asked to adjust the ITD of a separate broadband noise pointer—which alternated with the test stimulus—until the lateral positions of test and pointer matched.

⁸On the other hand, the conflicting cue (ILD) might simply impart a center bias to judgments of the main cue (ITD) in these studies. In that case, the cue conflict should not have prevented upweighting.

Abel, S. M., and Kunov, H. (1983). “Lateralization based on interaural phase differences: Effects of frequency, amplitude, duration, and shape of rise/decay,” *J. Acoust. Soc. Am.* **73**, 955–960.

Akeroyd, M. A., and Bernstein, L. R. (2001). “The variation across time of sensitivity to interaural disparities: Behavioral measurements and quantitative analyses,” *J. Acoust. Soc. Am.* **110**, 2516–2526.

Berg, B. G. (1989). “Analysis of weights in multiple observation tasks,” *J. Acoust. Soc. Am.* **86**, 1743–1746.

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.

Buell, T. N., Griffin, S. J., and Bernstein, L. R. (2008). “Listeners’ sensitivity to ‘onset/offset’ and ‘ongoing’ interaural delays in high-frequency, sinusoidally amplitude-modulated tones,” *J. Acoust. Soc. Am.* **123**, 279–294.

Buell, T. N., Trahiotis, C., and Bernstein, L. R. (1991). “Lateralization of low-frequency tones: Relative potency of gating and ongoing interaural delays,” *J. Acoust. Soc. Am.* **90**, 3077–3085.

Clifton, R. K., and Freyman, R. L. (1989). “Effect of click rate and delay on breakdown of the precedence effect,” *Percept. Psychophys.* **46**, 139–145.

Dizon, R. M., Culling, J. F., Litovsky, R. Y., Shinn-Cunningham, B. G., and Colburn, H. S. (1998). “On the development of a post-onset temporal weighting function,” *Assoc. Res. Otolaryngol. Abstr.* **21**, 42.

Franssen, N. V. (1962). *Sterophony* (Philips Technical Library, Eindhoven, The Netherlands).

Freyman, R. L., Zurek, P. M., Balakrishnan, Y., and Chiang, Y. C. (1997). “Onset dominance in lateralization,” *J. Acoust. Soc. Am.* **101**, 1649–1659.

Glanzer, M., and Cunitz, A. R. (1966). “Two storage mechanisms in free recall,” *J. Verbal Learn. Verbal Behav.* **5**, 351–360.

Grantham, D. W. (1984). “Discrimination of dynamic interaural intensity differences,” *J. Acoust. Soc. Am.* **76**, 71–76.

Grantham, D. W. (1997). “Auditory motion perception: Snapshots revis-

- ited," in *Binaural and Spatial Hearing in Real and Virtual Environments*, edited by R. H. Gilkey and T. R. Anderson (Lawrence Erlbaum Associates, Mahwah, NJ), pp. 295–313.
- Grantham, D. W., and Wightman, F. L. (1978). "Detectability of varying interaural temporal differences," *J. Acoust. Soc. Am.* **63**, 511–523.
- Hafter, 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 R. H. Gilkey and T. R. Anderson (Lawrence Erlbaum Associates, Mahwah, NJ), pp. 211–232.
- 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.
- Hartmann, W. M., and Rakerd, B. (1989). "Localization of sound in rooms IV: The Franssen effect," *J. Acoust. Soc. Am.* **86**, 1366–1373.
- Henning, G. B. (1974). "Detectability of interaural delay in high-frequency complex wave-forms," *J. Acoust. Soc. Am.* **55**, 84–90.
- Houtgast, T., and Aoki, S. (1994). "Stimulus-onset dominance in the perception of binaural information," *Hear. Res.* **72**, 29–36.
- Kunov, H., and Abel, S. M. (1981). "Effects of rise/decay time on the lateralization of interaurally delayed 1-kHz tones," *J. Acoust. Soc. Am.* **69**, 769–773.
- Lindemann, W. (1986). "Extension of a binaural cross-correlation model by contralateral inhibition. II. The law of the first wave front," *J. Acoust. Soc. Am.* **80**, 1623–1630.
- Lutfi, R. A. (1995). "Correlation coefficients and correlation ratios as estimates of observer weights in multiple-observation tasks," *J. Acoust. Soc. Am.* **97**, 1333–1334.
- Macpherson, E. A., and Wagner, M. L. (2008). "Temporal weighting of cues for vertical-plane sound localization," *Assoc. Res. Otolaryngol. Abstr.* **31**, 301.
- McFadden, D., and Pasanen, E. G. (1975). "Binaural beats at high frequencies," *Science* **190**, 394–396.
- McFadden, D., and Pasanen, E. G. (1976). "Lateralization of high frequencies based on interaural time differences," *J. Acoust. Soc. Am.* **59**, 634–639.
- Nuetzel, J. M., and Hafter, E. R. (1976). "Lateralization of complex wave-forms: Effects of fine structure, amplitude, and duration," *J. Acoust. Soc. Am.* **60**, 1339–1346.
- Rakerd, B., and Hartmann, W. M. (1985). "Localization of sound in rooms, II: The effects of a single reflecting surface," *J. Acoust. Soc. Am.* **78**, 524–533.
- Richards, V. M., and Zhu, S. P. (1994). "Relative estimates of combination weights, decision criteria, and internal noise based on correlation coefficients," *J. Acoust. Soc. Am.* **95**, 423–434.
- Saberi, K. (1996). "Observer weighting of interaural delays in filtered impulses," *Percept. Psychophys.* **58**, 1037–1046.
- Saberi, K., and Perrott, D. R. (1995). "Lateralization of click-trains with opposing onset and ongoing interaural delays," *Acustica* **81**, 272–275.
- Sadrulodabai, T., and Sorkin, R. D. (1999). "Effect of temporal position, proportional variance, and proportional duration on decision weights in temporal pattern discrimination," *J. Acoust. Soc. Am.* **105**, 358–365.
- Schroger, E. (1996). "Interaural time and level differences: Integrated or separated processing?," *Hear. Res.* **96**, 191–198.
- Shinn-Cunningham, B. G., Zurek, P. M., and Durlach, N. I. (1993). "Adjustment and discrimination measures of the precedence effect," *J. Acoust. Soc. Am.* **93**, 2923–2932.
- 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. (2000). "Observer weighting in sound localization," Ph.D. thesis, University of California, Berkeley, Berkeley, CA.
- Stecker, G. C., and Hafter, E. R. (2002). "Temporal weighting in sound localization," *J. Acoust. Soc. Am.* **112**, 1046–1057.
- Stecker, G. C., Harrington, I. A., and Middlebrooks, J. C. (2005). "Location coding by opponent neural populations in the auditory cortex," *PLoS Biol.* **3**, e78.
- Stecker, G. C., and Middlebrooks, J. C. (2003). "Distributed coding of sound locations in the auditory cortex," *Biol. Cybern.* **89**, 341–349.
- Stellmack, M. A., Dye, R. H., and Guzman, S. J. (1999). "Observer weighting of interaural delays in source and echo clicks," *J. Acoust. Soc. Am.* **105**, 377–387.
- Strutt, J. W. (1907). "On our perception of sound direction," *Philos. Mag.* **13**, 214–232.
- Tardif, E., Murray, M. M., Meylan, R., Spierer, L., and Clarke, S. (2006). "The spatio-temporal brain dynamics of processing and integrating sound localization cues in humans," *Brain Res.* **1092**, 161–176.
- Tobias, J. V., and Schubert, E. R. (1959). "Effective onset duration of auditory stimuli," *J. Acoust. Soc. Am.* **31**, 1595–1605.
- Ungan, P., Yagcioglu, S., and Goksoy, C. (2001). "Differences between the n1 waves of the responses to interaural time and intensity disparities: Scalp topography and dipole sources," *Clin. Neurophysiol.* **112**, 485–498.
- 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.
- Wallach, H., Newman, E. B., and Rosenzweig, M. R. (1949). "The precedence effect in sound localization," *Am. J. Psychol.* **62**, 315–336.
- Yamada, K., Kaga, K., Uno, A., and Shindo, M. (1996). "Sound lateralization in patients with lesions including the auditory cortex: Comparison of interaural time difference (ITD) discrimination and interaural intensity difference (IID) discrimination," *Hear. Res.* **101**, 173–180.
- Zurek, P. M. (1980). "The precedence effect and its possible role in the avoidance of inter-aural ambiguities," *J. Acoust. Soc. Am.* **67**, 952–964.