




Spectral weighting functions for localization of complex sound. II. The effect of competing noise^{a)}

Monica L. Folkerts,^{1,b)}  Erin M. Picou,²  and G. Christopher Stecker³ 

¹Department of Otolaryngology–Head and Neck Surgery, University of North Carolina at Chapel Hill, 170 Manning Drive, Chapel Hill, North Carolina 27599-7070, USA

²Department of Hearing and Speech Sciences, Vanderbilt University Medical Center, 1215 21st Avenue South, Nashville, Tennessee 37232, USA

³Center for Hearing Research, Boys Town National Research Hospital, 555 North 30th Street, Omaha, Nebraska 68131, USA

ABSTRACT:

Spectral weighting of sound localization cues was measured in the presence of three levels of competing noise presented in the free field. Target stimuli were complex tones containing seven tonal components, presented from an $\sim 120^\circ$ range of frontal azimuths. Competitors were two independent Gaussian noises presented from 90° left and right azimuth at one of three levels yielding +9, 0, and -6 dB signal-to-noise ratio. Results revealed the greatest perceptual weight for components within the interaural time difference (ITD) “dominance region,” which was found previously to peak around the 800-Hz component in quiet [Folkerts and Stecker (2022) *J. Acoust. Soc. Am.* **151**, 3409–3425]. Here, peak weights were shifted toward lower-frequency components (i.e., 400 Hz) in all competing noise conditions. These results contradict the hypothesis of a shift in the peak weights toward higher frequencies based on previous behavioral localization performance in competing noise but are consistent with binaural cue sensitivity, availability, and reliability; measured low-frequency ITD cues within the dominance region were least disrupted by the presence of competing noise. © 2023 Acoustical Society of America.

<https://doi.org/10.1121/10.0020294>

(Received 7 March 2023; revised 29 June 2023; accepted 10 July 2023; published online 25 July 2023)

[Editor: Pavel Zahorik]

Pages: 494–501

I. INTRODUCTION

The localization of real sound sources requires the integration of multiple spatially informative acoustic cues, including interaural time differences (ITDs) and interaural level differences (ILDs), which are themselves distributed across multiple frequency components of complex sounds. For narrowband stimuli, such as pure tones, acoustic effects limit the effectiveness of each cue at various frequencies. For example, the size and shape of the head implies ILDs are more informative above 1000 Hz, whereas listeners become insensitive to ITD cues for tone frequencies above approximately 1500 Hz (Blauert and Allen, 1997; Brughera *et al.*, 2013). Thus, the localization of narrowband sounds relies primarily on ITD cues at low frequencies and ILD cues at high frequencies (Strutt, 1907). Over the years, that view—labeled the “duplex theory” of sound localization—has been amended to consider the role of high-frequency ITD in modulated sounds (e.g., Henning, 1974), the sensitivity of listeners to ILD across a broad range of frequencies (e.g., Yost, 1981), and the availability of large low-frequency ILD for nearby sources (Brungart *et al.*, 1999).

Macpherson and Middlebrooks (2002; see also Wightman and Kistler 1992) considered the relative contributions of low- and high-frequency cues in the (virtual) localization of filtered and broadband sounds. Consistent with the duplex theory, localization of low-pass stimuli was strongly affected by synthetic manipulation of the ITD cue, and high-pass localization was most affected by manipulation of the ILD. Broadband localization was affected by both cues; however, ITD manipulation had a greater effect, suggesting that broadband sound localization relies predominantly on the availability of low-frequency ITD cues.

More recently, observer-weighting methods have been used to directly estimate the contributions of various frequency components to the localization and lateralization of complex sounds presented in quiet (Folkerts and Stecker, 2022). Those methods introduce small random variations in the spatial cues of each component, then use statistical techniques (e.g., multiple regression) to relate those variations to listeners’ spatial judgments of the overall stimulus. The influence of each component is expressed as a perceptual “weight” derived from the regression coefficients (Ahumada and Lovell, 1971).

Spectral weighting functions (SWFs) display these weights as a function of frequency. Results of prior SWF measurements provide insight into the frequency-specific dominance for ITD and ILD. For example, Ahrens *et al.* (2020) measured lateralization SWFs for a broadband stimulus

^{a)}Portions of this work were presented in “Spectral weighting in sound localization: Effects of competing noise,” 181th Meeting of the Acoustical Society of America, Seattle, WA, USA, 29 November–3 December 2021, and dissertation by M.L.F.

^{b)}Electronic mail: monica@spatialhearing.org

comprised of 11 bands of noise, each 1-ERB (equivalent rectangular bandwidth; Moore and Glasberg, 1983) wide and with center frequencies ranging from 442 to 5544 Hz (a 1-ERB gap separated adjacent bands) and carrying either ITD or ILD (the other cue was set to zero). ITD-based lateralization produced SWFs with elevated weight on the 442-Hz centered band while ILD-based lateralization produced SWFs peaking around the 5544-Hz band. For broadband stimuli with ITDs and ILDs available, the mid-frequencies are particularly important. For example, for broadband stimuli (100–6400 Hz) with tonal or noise components carrying ITD and ILD, Folkerts and Stecker (2022) found that lateralization SWFs exhibited the greatest weight on the 800-Hz ITD component. During free-field localization, SWFs closely matched the ITD weights in lateralization SWFs with the greatest weight on the 800-Hz component, suggesting that broadband localization relied mainly on the ITD cue in this region (Folkerts and Stecker, 2022; see also Macpherson and Middlebrooks, 2002).

A SWF peak around 800 Hz is also consistent with the ITD “dominance region”—wherein cue salience peaks around 600–800 Hz—described in older studies on the lateralization of filtered clicks (Bilsen and Raatgever, 1973; Tollin and Henning, 1999). That observation is thought to reflect greater reliability of the ITD cue in that frequency region in natural listening environments (i.e., rooms) where reverberation and other acoustic effects may occur. This seems to be borne out in the impacts of reverberation on SWFs (Folkerts and Stecker, 2022), where reverberation increased the relative weight of components around 800 Hz compared to anechoic conditions.

Another element common in real-world environments is competing noise. SWFs have not been measured in the presence of competing noise, and they could be impacted by masking, interference, and/or localization biases induced by competing noise. Based on existing evidence of behavioral localization performance in competing noise, it is expected that SWFs in the presence of competing noise would demonstrate increased weight of high-frequency components. For instance, evidence suggests that broadband competing sounds are more disruptive for low-frequency target signals than for high-frequency or broadband target signals. This effect has been demonstrated for competing sounds presented from frontal or lateral angles (Lorenzi *et al.*, 1999), behind the listener (Abel and Hay, 1996), or from random locations (Brungart and Simpson, 2009). Thus, these results suggest that high-frequency cues are important for precise localization in the presence of competing noise and, in turn, relatively higher frequencies might be emphasized in such conditions. That is, the prediction is that for a broadband signal presented in competing noise, perceptual weights would increase for high-frequency components and decrease for components within the dominance region.

An alternative hypothesis is that competing noise could *increase* the relative importance of lower-frequency cues contained within the dominance region under the assumption that greater ITD sensitivity at those frequencies is

accompanied by robust coding in noise. Consistent with that view, Yost *et al.* (1971) and Yost (1975) reported greater effects of competing noise on ITD sensitivity for high-frequency targets (filtered clicks, bands of noise, and modulated tones) than for low-frequency (below 1500 Hz) targets. Similarly in the sound field, Jacobsen (1976) reported greater effects of competing noise on the minimum audible angle (MAA; the just noticeable difference in azimuth) for tonal targets at 2000 versus 500 Hz. Whereas MAA for higher-frequency targets increased from 3° in quiet to 6° in noise, MAA for the low-frequency tone was unaffected by noise. Here, we used binaural cue analysis (see the supplementary material¹) to measure the cues available for 400- and 800-Hz tones (ITD) and 3200- and 6400-Hz tones (ILD) in the presence of competing noise. Analysis revealed an azimuth-dependent increase in ITD variability (supplementary Fig. 1) and a decrease in interaural correlation (IAC; supplementary Fig. 2) with increasing noise. The effect was greater for the 800- than for the 400-Hz tone. ILD values for the high-frequency tones decreased with increasing noise (supplementary Fig. 3). Together, these observations suggest greater robustness to noise for lower-frequency components within the ITD dominance region. Consequently, competing noise may lead to greater relative weight on these low-frequency components compared to quiet conditions.

The purpose of the current study was to evaluate the effect of competing noise on spectral dominance in free-field localization. Specifically, SWFs were measured during the localization of complex tones in the presence of competing noise. The target complex tones were comprised of tones within the ITD dominance region and across a range of high frequencies.

II. METHODS

SWFs were measured in the presence of competing noise for participants with normal hearing. Two stimulus conditions from Folkerts and Stecker (2022) containing high frequencies (denoted T1 and T3 in that study) were used. SWFs were measured (and analyzed) in a similar manner to Folkerts and Stecker (2022) with the exception that Gaussian noise was presented in three signal-to-noise ratio (SNR) conditions (+9, 0, and –6 dB) for each stimulus condition. SWFs in competing noise were compared to established SWFs measured in quiet in a nonoverlapping group of participants (from Folkerts and Stecker, 2022).

A. Participants

Ten adults (eight females) aged 21–28 years old were recruited from the Vanderbilt University community. One participant (0515) was M.L.F. Normal hearing was confirmed via pure-tone audiometry. Thresholds were better than 20 dB HL and differed less than 15 dB between left and right ears at all octave frequencies from 250 to 8000 Hz. Participants were monetarily compensated. Approval was obtained for experimental procedures from Vanderbilt

University Medical Center Institutional Review Board (IRB No. 191952).

B. Stimuli

All stimuli were generated in MATLAB (The Mathworks, Natick, MA) and synthesized at 48 kHz. Stimuli were presented through a Dante audio-over-ethernet network (Focusrite Rednet, El Segundo, CA) with digital amplification (Ashly ne820PE, Webster, NY) for playback to a 64-loudspeaker array (Meyer MM-4, Berkeley, CA) in the Vanderbilt Bill Wilkerson Anechoic Chamber Laboratory (ACL; $4.6 \times 6.4 \times 6.7$ m; Eckel Industries, Cambridge, MA). The loudspeakers in the ACL are positioned at ear height, spanning 360° (5.625° of separation) with a 2-m radius.

Stimuli included tone complexes serving as localization targets (used to measure SWFs) and Gaussian noises serving as competitors. Each target signal consisted of seven pure-tone components presented at equal amplitude. The duration of the tone complex was 100 ms, including 10-ms \cos^2 onset/offset ramps. Component frequencies were configured either as octave frequencies from 100 to 6400 Hz (for stimulus condition T1) or as harmonics of 800 Hz spanning 800–5600 Hz (stimulus condition T3). The overall level of the target signal was held constant at 60 dB sound pressure level (SPL) across SNR conditions. The spatial configuration of each of the seven frequency components was manipulated from trial to trial to introduce spatial variation (“jitter”) suitable for measuring perceptual weights in the same manner as in Folkerts and Stecker (2022). On each trial, a “base” azimuth was chosen from 11 possible locations: -56.26° to $+56.26^\circ$ in 11.25° steps. Five adjacent loudspeakers centered on the base azimuth [-11.25° , -5.625° , 0° , $+5.625^\circ$, and $+11.25^\circ$ (relative to the base)] constituted the set of source loudspeakers from which individual components of the stimulus were presented on a given trial. Each component was randomly and independently assigned to one of the five source loudspeakers.

The competing noises were two independent 200-ms samples of Gaussian white noise. These were presented simultaneously from the left and right lateral loudspeakers (-90° and $+90^\circ$), beginning 50 ms before the target signal. The level of the competing noise was dependent on the SNR condition: -6 dB SNR (66 dB SPL), 0 dB SNR (60 dB SPL), and $+9$ dB SNR (51 dB SPL).

C. Procedure

A touch-sensitive display (Apple iPad Air, Cupertino, CA) was mounted at a comfortable distance (~ 0.5 m in front and ~ 0.5 m below ear level) from the participant. This was used to record localization responses aligned to a top-down visual schematic of the room and loudspeaker array displayed on the screen (Fig. 1). On each trial, participants were instructed to sit upright and face directly forward (toward the loudspeaker at 0°) before and during each stimulus presentation. These instructions helped to ensure

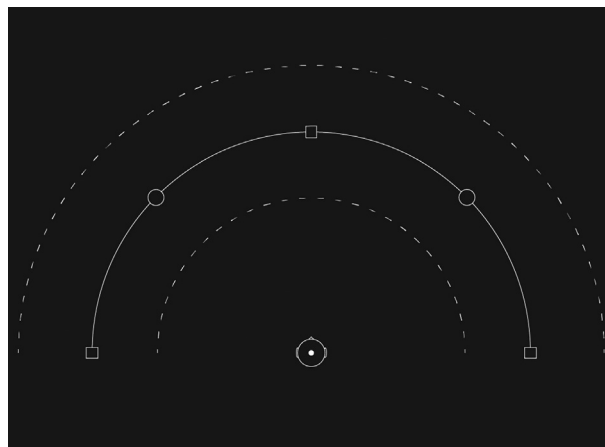


FIG. 1. Schematic used to record localization responses. The head marks the participant with a centered white dot at the middle bottom portion of the 180° arc. The solid white line of the arc indicates the loudspeaker array. Landmarks are included along the solid line, indicating 0° (top open square), -90° and $+90^\circ$ (open bottom squares), and -45° and $+45^\circ$ (open middle circles). In the sound field, -45° and $+45^\circ$ were marked by unused circular loudspeakers. Dashed lines were used to indicate the distance (if necessary), either in front or behind the loudspeaker array.

participants received expected spatial cues. Immediately following each single presentation of the stimulus, participants were instructed to make an eye movement to the target signal’s perceived location and then record that location on the schematic diagram by touching the iPad screen. Participants were instructed that on any trial in which the lateral percept appeared “wide” or “split,” they should indicate the leftmost edge or leftmost perceived auditory image. Because frequency components have an equal probability of being presented from the leftmost loudspeaker (-11.25°) relative to the base loudspeaker, this instruction helps ensure that SWFs remain unbiased should these percepts occur. Debriefing participants revealed no reports of such percepts occurring for the stimuli used here. That is, as in previous studies (e.g., Stecker *et al.*, 2013), stimuli were reliably perceived as spatially fused and localizable. Thus, weights are valid indicators of components’ relative contributions to the fused localization judgment. The response azimuth was computed from the touch screen response and recorded as the localization judgment.

The base loudspeaker was selected pseudorandomly across trials with each base value presented 6 times per run of 66 trials. Participants completed 6 runs (396 trials) for each of the 6 conditions [i.e., 2 stimulus conditions (T1 and T3) and 3 SNR conditions (-6 , 0, and $+9$ dB)].

D. Analysis

SWFs were calculated separately for each participant and condition. Within each 66-trial run, the localization judgment was normalized by rank transform across each 66-trial run. Rank normalization ensures a uniform distribution of response data that is appropriate for linear regression, reducing or eliminating effects of across-listener differences in response bias, range, and linearity (see Stecker *et al.*,

2013). Perceptual weights for each of the seven frequency components were estimated by multiple linear regression of the rank-transformed responses, θ_R , onto the azimuth values of each component, θ_i :

$$\theta_R = \sum_{i=1}^7 \beta_i \theta_i + k. \tag{1}$$

Weights were computed by normalizing regression coefficients, β_i , such that absolute values summed to one across weights:

$$w_i = \beta_i / \sum_{j=1}^7 |\beta_j|. \tag{2}$$

The normalized weights, w_i , indicate the relative influence of the frequency component, i , on participants' localization judgments. Plotted together, the normalized weights constitute the SWF for each participant in each condition. Group average SWFs were calculated by taking the arithmetic mean normalized weight across participants for each component.

Folkerts and Stecker (2022) computed the “average ratio” (AR) as a univariate measure of nonuniformity, specifically with an emphasis on the 800-Hz component in the ITD dominance region (AR_{800}). In the current study, the AR_{800} was used to interpret the prominence of the 800-Hz component across quiet and SNR conditions. The AR_{800} was defined as the ratio of weight on the 800-Hz component to the mean of other weights,

$$AR_{800} = w_{800\text{Hz}} / \left(\sum_{j \neq 800\text{Hz}} w_j / 6 \right). \tag{3}$$

The AR_{800} was calculated for each participant in each condition, and the group average was the arithmetic mean of the AR_{800} across participants.

As in Folkerts and Stecker (2022), the current study computed SWF confidence intervals and evaluated planned comparisons of weight and AR_{800} by nonparametric bootstrap tests. Bootstrapped confidence intervals on mean weight values were computed by resampling weights with replacement across subjects to generate 2000 bootstrap replicates. The mean weight was computed for each replicate to estimate the sampling distribution of mean weights. Confidence intervals were computed at the 95% level by taking the 2.5 and 97.5 percentile points from this sampling distribution.

Null-hypothesis significance tests used a similar approach. Each measure was resampled across participants to generate 2000 bootstrap replicates. A statistic of interest (e.g., mean or difference between two means) was then computed for each bootstrap replicate to estimate the corresponding sampling distribution. The proportion of bootstrap replicates falling at or below the null-hypothesis prediction (e.g., $AR \leq 1$) defined the (one-sided) p -value, which is

expressed to one significant digit. For two-sided statistical tests, the p -value was computed as the minimum of proportions falling on either side of the prediction, doubled. When any proportion was zero (i.e., the bootstrap sampling distribution did not overlap the null-hypothesis prediction), p -values are listed at the resolution of the bootstrap test itself (i.e., $p < 0.0005$ for 2000-fold bootstrap).

III. RESULTS AND DISCUSSION

A. SWFs

Figure 2 displays the mean SWF obtained in each of the six conditions (i.e., two stimulus conditions in three competing noise conditions). Included for comparison are SWFs measured in quiet (leftmost column) for the two stimulus conditions, previously reported for a different group of participants in Folkerts and Stecker (2022). For ease of comparison, the SWFs measured in quiet are also plotted as thin lines on each competing noise panel. Significant differences between component weights in quiet versus competing noise are indicated with asterisks. In the presence of various levels of noise, the overall shape of the SWFs generally appears consistent with SWFs found in quiet. There were no significant differences between the weight placed on the 800-Hz component in quiet and in the presence of competing noise. However, in condition T1 (octave tones 100–6400 Hz), the 400-Hz component received a greater weight in the –6 dB SNR condition ($p < 0.05$) and +9 dB SNR condition ($p < 0.001$) than in quiet, resulting in a slight “widening” of the ITD dominance-region peak toward lower frequencies. These results are contrary to the hypothesized shift in weight maxima toward higher frequencies (Abel and Hay, 1996; Lorenzi *et al.*, 1999; Brungart and Simpson, 2009). However, the shift toward lower frequencies is consistent with the alternative hypothesis of increased weights for lower-frequency components within the dominance region. Across most conditions, the highest frequency components in both stimulus conditions (6400 Hz and 5600 Hz for T1 and T3, respectively) were weighted lower in competing noise than in the quiet ($p < 0.02$ across all SNRs).

It is informative to consider the differences in stimuli of the current study and those used to derive the hypothesis of increased weights for higher-frequency components in competing noise relative to weights in quiet (Abel and Hay, 1996; Lorenzi *et al.*, 1999; Brungart and Simpson, 2009). Target stimuli in the current study were 100 ms complex non-modulated tones, whereas others presented longer-duration stimuli with ongoing envelope fluctuations. Abel and Hay (1996) presented 300 ms 1/3-octave noise bands centered at 500 or 4000 Hz, Lorenzi *et al.* (1999) presented 300-ms pulse trains at a rate of 100 Hz, and Brungart and Simpson (2009) presented 205-ms trains of white noise bursts at a rate of 22 Hz. It is possible that the available binaural cues differ between tones and modulated sounds, particularly as a consequence of envelope fluctuations that enhance binaural sensitivity (Stecker *et al.*, 2021) but may be reduced in the presence of continuous masking noise.

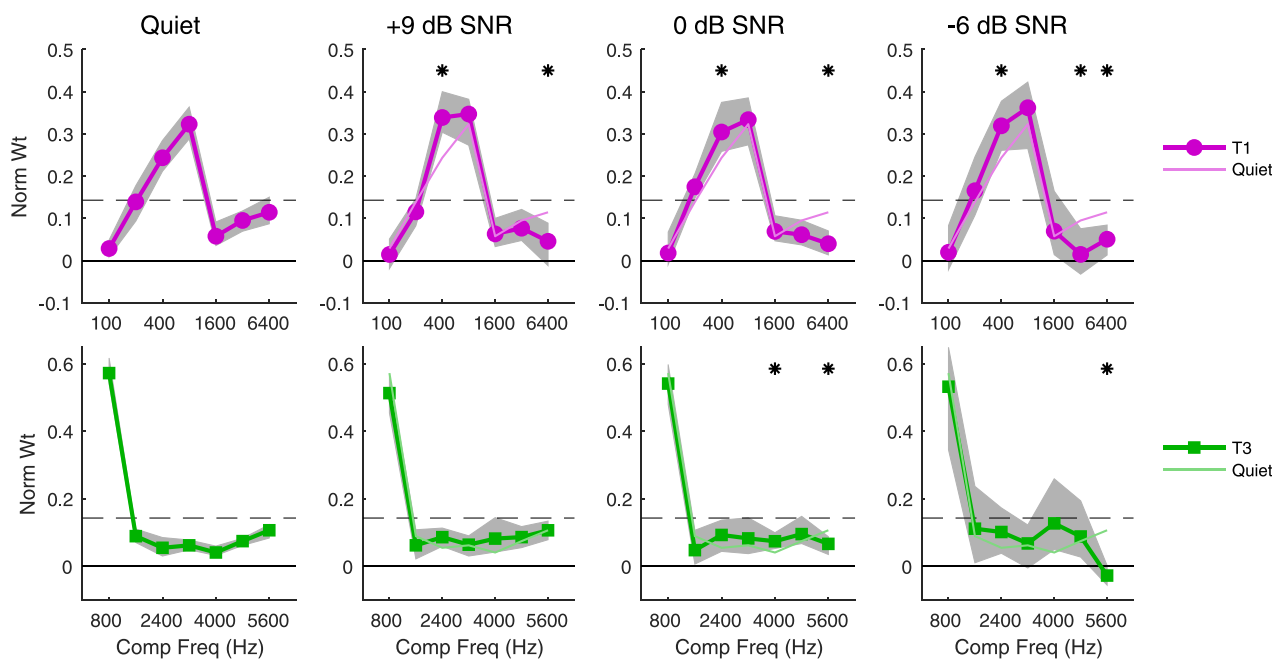


FIG. 2. (Color online) Mean SWFs obtained in various levels of competing noise (columns, +9 dB SNR, 0 dB SNR, and -6 dB SNR), including no noise data from Folkerts and Stecker (2022; leftmost column, “quiet”). Top and bottom panels plot SWF data for stimulus conditions T1 (octave tones, 100–6400 Hz, magenta circles) and T3 (harmonic tones, 800–5600 Hz, green squares), respectively. Symbols and thick lines plot cross-participant mean normalized weight as a function of component frequency. Shaded regions indicate bootstrapped $\pm 95\%$ confidence intervals on each mean weight. Dashed lines indicate the expected value (1/7) for uniform weighting across components. Quiet SWFs are replotted as thin colored lines in competing noise panels for purposes of comparison. Bootstrapped, two-tailed, significant differences ($p < 0.05$) between weights in the quiet and competing noise weights are indicated with asterisks (*) at the top of each panel.

The timing and spatial configuration of competing noise also differed between studies. Simpson *et al.* (2011) found a monotonically decreasing impact of competing noise on localization accuracy as stimulus onset asynchrony (SOA, the delay of target onset relative to noise onset) increased from 0 to 240 ms. The current study employed a SOA of 50 ms compared to 25 ms in Brungart and Simpson (2009) and 300 ms in Lorenzi *et al.* (1999). Abel and Hay (1996) presented noise continuously (infinite SOA). Thus, relative to the current study, stimulus timing would predict opposite effects across the other studies and seems unlikely to explain the observed differences. In contrast, the geometry of noise sources used in the current study (two lateral sources at $\pm 90^\circ$ azimuth) might have impacted high-frequency ILD cues more strongly than in the other studies. Abel and Hay (1996) presented three simultaneous rearward noise sources, whereas Lorenzi *et al.* (1999) presented only a single source at -90° , 0° , or $+90^\circ$ azimuth in separate conditions. Brungart and Simpson (2009) also presented a single source with random azimuth and elevation. The choice here to present noise sources laterally and symmetrically was made to emulate real-world environments, where uncorrelated noises arise from various locations, and maximize localization errors (Good *et al.*, 1997; Lorenzi *et al.*, 1999). Symmetrical placement of competing noise reduced the ILD cue, especially at high frequencies (supplementary Fig. 2),¹ whereas a single competing noise source might have left high-frequency ILD cues more intact in the studies by Lorenzi *et al.* (1999) and Brungart and Simpson (2009).

Individual SWFs for each participant, including SWFs measured in a nonoverlapping group of participants reported in quiet by Folkerts and Stecker (2022), are displayed in Fig. 3. The overall trend of SWF shapes is consistent with the results found in the mean data. That is, a consistent weight maximum in the ITD dominance region (the 800-Hz component), with a possible shift of the peak region toward lower frequencies [i.e., 400 Hz; stimulus condition T1; Fig. 3 (top)].

Compared to individual SWFs in quiet, the SWFs in competing noise progressively become more variable across participants, especially as the SNR unfavorably decreased. This is also evident in the widening of confidence intervals with decreasing SNR in Fig. 2.

AR₈₀₀ distributions are displayed in Fig. 4 as violin plots (Hintze and Nelson, 1998; Bechtold, 2016), including the AR₈₀₀ distributions in the same target stimulus conditions in quiet (Folkerts and Stecker, 2022) for comparison purposes. The AR₈₀₀ was used in the current analysis for comparison to quiet data, where AR₈₀₀ focused on the observed dominance-region peak at 800 Hz. All AR₈₀₀ values across SNR conditions were significantly greater than one, indicated by asterisks. There were no significant differences in AR₈₀₀ across quiet and SNR conditions. The consistent nonuniformity across SNR conditions indicates that participants continue to use the cues in the ITD dominance region, even in the presence of competing noise. However, this measurement may not capture the “broadening” of the peak in stimulus condition T1 (octave tones 100–6400 Hz),

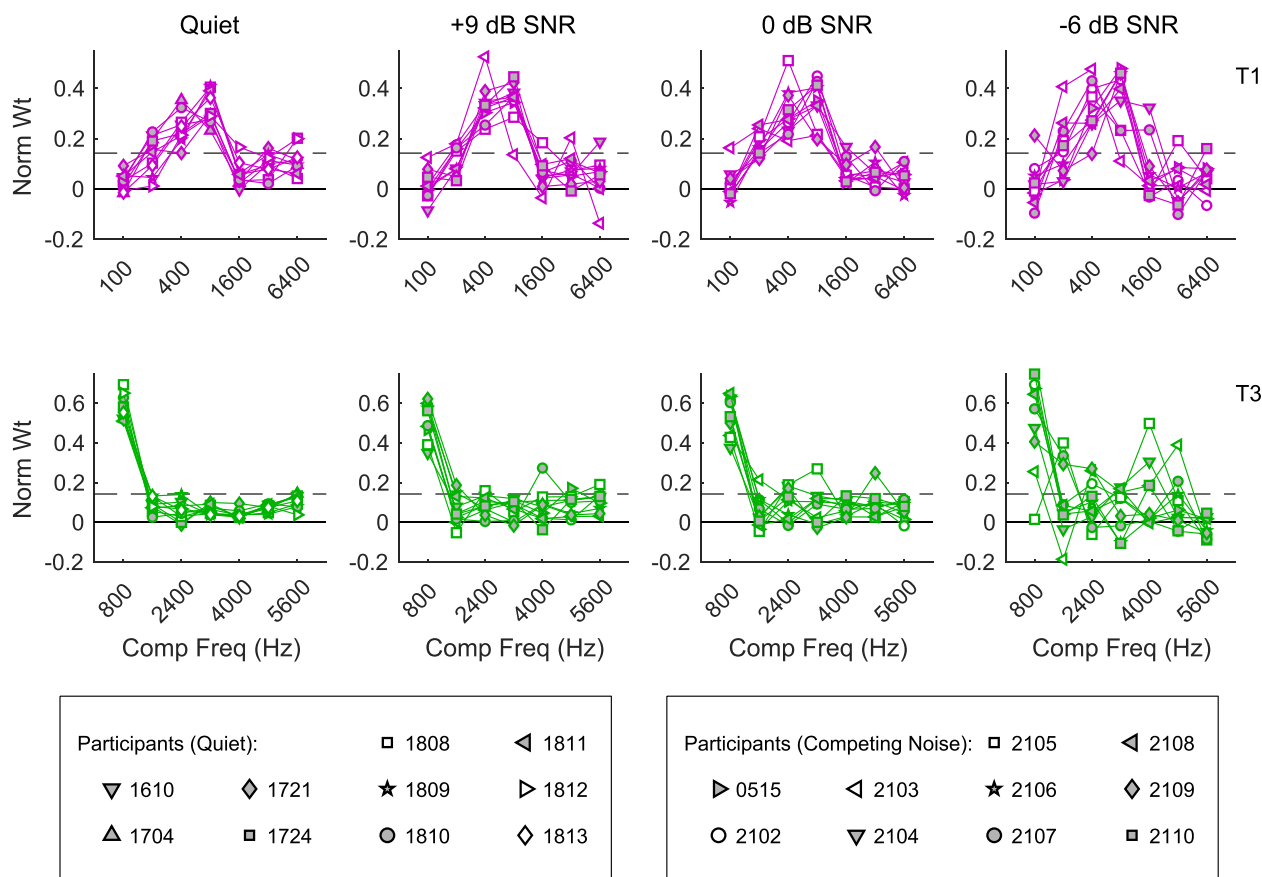


FIG. 3. (Color online) Individual-participant SWFs found in various levels of competing noise (columns, +9 dB SNR, 0 dB SNR, and -6 dB SNR), including no noise data from Folkerts and Stecker (2022; leftmost column, “quiet”). Panels are arranged identically to those in Fig. 1, plot SWFs obtained for individual participants (symbols; legend at bottom). Note that one group of participants completed the competing noise conditions; a nonoverlapping group of participants completed the quiet experiment.

which seemed to occur as a significant increase in the 400-Hz component for all SNR conditions.

The SWFs in all SNR conditions reveal a decrease in the relative weight for high-frequency components, contrary to the expectation that the maxima of 800-Hz component found in quiet would shift toward high-frequency components in the presence of competing noise. The largest weights were found to remain within the ITD dominance region (with a slight expansion of the peak toward lower frequencies), even in the least favorable SNR conditions (-6 dB). SWFs continue to reveal elevated weights for components within the dominance region. The data support a persistent ITD dominance region in the presence of competing noise. The down-weighting of the high-frequency cues is inconsistent with previous behavioral studies, which suggested high-frequency information to be particularly important for localization in noise (Abel and Hay, 1996; Lorenzi et al., 1999; Brungart and Simpson, 2009). The current results are, however, consistent with the alternative hypothesis derived from binaural cue sensitivity and analysis, where the binaural cues for the low-frequency portion of the dominance region are less disrupted by competing noise (as presented here, laterally and symmetrically) than binaural cues for higher frequencies.

B. Localization performance in competing noise

The shapes of SWFs across SNR conditions are similar; however, the variability in weighting across participants increases as the SNR decreases. This variability is evident in Fig. 2, where the confidence interval (gray shaded region) widens in the -6 dB SNR condition, in Fig. 3, where the

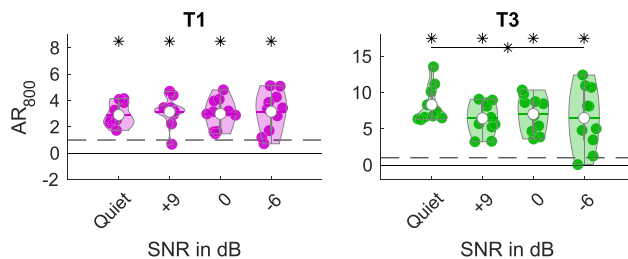


FIG. 4. (Color online) Violin plots (Hintze and Nelson, 1998; Bechtold, 2016) of AR_{800} values (vertical axis) are displayed for each stimulus condition (panels), plotted across SNR conditions (-6, 0, and +9 dB SNR; horizontal axis), including in quiet. Adapted from Folkerts and Stecker (2022). Colored circles in each panel plot AR_{800} for individual participants; violin plots indicate the density (width of each violin) and mean (white circle and colored line) of obtained values. Dashed lines indicate the expectation for uniform spectral weighting $AR=1$. Asterisks (*) indicate conditions in which AR_{800} significantly exceeded this value ($AR_{800} > 1$, $p < 0.01$ by 2000-fold bootstrap test).

variability of SWF shapes across participants increases in the -6 dB SNR condition, and in Fig. 4, where the range of individual AR_{800} values (colored circles) increases. A reasonable explanation may be that localization performance degrades in the presence of competing noise, especially as the SNR becomes less favorable (Good *et al.*, 1997; Lorenzi *et al.*, 1999; Brungart and Simpson, 2009). To explore this possibility, localization performance was examined for the current data by measuring the root mean square (RMS) error statistic and the regression statistic, R^2 .

Recall that to measure SWFs, spatial jitter was introduced so that the multiple linear regression model [Eq. (1)] could be used. Thus, the individual components did not match exactly in azimuth, and there is no objectively “correct” localization response on any given trial. Instead, accurate localization was defined as a participant’s response falling within the 22.5° range of five adjacent loudspeakers active on that trial. Localization responses outside of that range were classified as errors.

The RMS error statistic was used by Good and Gilkey (1996) and Lorenzi *et al.* (1999) to calculate a participant’s average error in degrees during localization in noise. RMS error accounts for variance and bias in localization responses. Here, RMS error was computed from the magnitude by which response azimuth exceeded the 22.5° range of spatial jitter (i.e., responses falling within the window were coded as zero error). The left panel of Fig. 5 displays the mean RMS error (across participants) in degrees as a function of SNR condition for each frequency stimulus. Bootstrapped difference tests at each SNR value were calculated between stimulus conditions T1 and T3 and between SNR conditions (within stimulus conditions). The RMS error in both stimulus conditions (T1 and T3) are relatively similar at $+9$ and 0 dB SNR as opposed to -6 dB SNR, where stimulus condition T3 has a higher degree of error than T1. However, the difference between T1 and T3 at the low SNR did not reach significance. For both stimulus conditions (T1 and T3), the RMS error was significantly larger in the -6 dB SNR condition than in the $+9$ dB and 0 dB conditions (lines with asterisks).

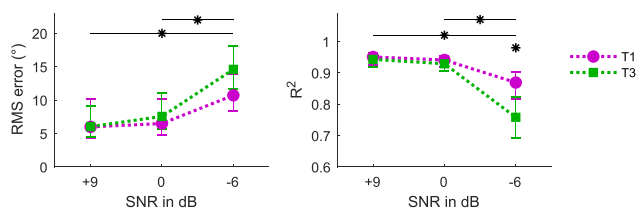


FIG. 5. (Color online) The left panel plots mean RMS error (in deg) across participants as a function of SNR in dB for stimulus conditions T1 (magenta circles; octave tones 100–6400 Hz) and T3 (green squares; harmonic tones 800–5600 Hz; legend to the right of panels). The right panel plots the mean R^2 value across participants as a function of SNR in dB for stimulus conditions T1 and T3. Error bars indicate bootstrapped $\pm 95\%$ confidence intervals on each mean value (RMS error or R^2). Asterisks (*) indicate bootstrapped, two-tailed, significant differences ($p < 0.05$) between (RMS error or R^2) values in stimulus condition T1 versus T3. Asterisks with lines indicate significant differences across SNR conditions (for both stimulus conditions and calculated separately).

Another approach to quantify localization error is to calculate the proportion of response variance accounted for by stimulus location variance, most commonly expressed by the regression statistic, R^2 . Good and Gilkey (1996) and Lorenzi *et al.* (1999) measured R^2 using regression with target location (i.e., the physically correct location) as a single predictor variable. That approach is precluded here because each trial presents components from a mixture of five adjacent loudspeakers. However, the multiple linear regression model used for weight calculation [Eq. (1)] captures the effects of stimulus location variance across all components. The R^2 statistic, thus, expresses the proportion of localization response variance, which is explainable by the linear model, i.e., the non-error variance. After obtaining each SWF measurement using the multiple linear regression model (for each participant in each condition), R^2 was generated by the model statistics. The right panel of Fig. 5 displays the mean R^2 value across participants as a function of SNR condition for each stimulus condition. Bootstrapped difference tests at each SNR value were calculated between stimulus conditions T1 and T3 and between SNR conditions (within stimulus conditions). Similar to the differences in RMS error across stimulus conditions, the R^2 values were similar for T1 and T3 at $+9$ and 0 dB SNR. At -6 dB SNR, the R^2 values were significantly different between stimulus conditions T1 and T3 (asterisk). For both stimulus conditions (T1 and T3), the R^2 value was significantly smaller in the -6 dB SNR condition than the $+9$ dB and 0 dB conditions (lines with asterisks).

Across both stimulus conditions, T1 (octave tones 100–6400 Hz) and T3 (harmonic tones 800–5600 Hz), the RMS errors and R^2 values display an increase in localization response error when the SNR is below 0 dB. These data are consistent with the RMS errors and R^2 values reported by Lorenzi *et al.* (1999) when using broadband and high-pass click stimuli. However, Lorenzi *et al.* (1999) found that response error rates increased for low-pass stimuli (1600 Hz cutoff) when compared to broadband and high-pass click stimuli. From this, Lorenzi *et al.* (1999) concluded that in the presence of a lateral masker, participants localize signals using high-frequency cues such as ILD or envelope ITD. In the current experiment, localization errors were larger for stimuli that encompassed mostly high frequencies (T3; harmonic tones 800–5600 Hz) than for stimuli that were broadband (T1; octave tones 100–6400 Hz). These results do not align with those of Lorenzi *et al.* (1999), who found no difference in localization performance between broadband and high-pass stimuli. However, there are differences in cue reliability (clicks versus tone complexes), competing noise configuration (one lateral noise versus two lateral noises), and localization error statistic derivation (see above).

IV. SUMMARY AND CONCLUSIONS

The purpose of the current study was to explore the effects of competing noise on SWFs. The SWFs of the current participants were compared to participants tested in quiet (Folkerts and Stecker, 2022). The results of this study revealed two main findings:

- (1) SWFs measured in the presence of competing noise revealed the greatest weight on the 800-Hz component of complex tones across all SNRs (+9, 0, and -6 dB). Compared to SWFs measured in quiet, the 400-Hz component revealed an elevated weight, and the highest frequency components revealed a reduced weight. These results are consistent with frequency-dependent ITD sensitivity in the presence of competing noise; and
- (2) although no systematic differences in the salience of the dominance region were found across SNRs, localization performance, and, consequentially, the performance of the model used to measure SWFs degraded when the SNR was -6 dB SNR. The degradation was more apparent for stimulus condition T3, which contained primarily high frequencies (i.e., 800–5600 Hz), than for stimulus condition T1, which contained lower frequencies (i.e., 400 Hz) with less disrupted ITD cues in the presence of competing noise. Therefore, the current results do not align with earlier work that noise least affects binaural processing of high-frequency stimuli (e.g., Lorenzi *et al.*, 1999).

ACKNOWLEDGMENTS

This work was supported by Grant Nos. R01 DC011548 and R01 DC016643 (G.C.S.) from the National Institute on Deafness and Other Communication Disorders (NIDCD). Additional support was provided by Vanderbilt University Medical Center. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIDCD or the National Institutes of Health. Additional thanks go to dissertation committee members Marc Brennan, Rene Gifford, and Todd Ricketts, and anonymous reviewers.

¹See supplementary material at <https://doi.org/10.1121/10.0020294> for supplementary text and figures regarding binaural recordings.

- Abel, S. M., and Hay, V. H. (1996). "Sound localization the interaction of aging, hearing loss and hearing protection," *Scand. Aud.* **25**, 3–12.
- Ahrens, A., Joshi, S. N., and Epp, B. (2020). "Perceptual weighting of binaural lateralization cues across frequency bands," *J. Assoc. Res. Otolaryngol.* **21**, 485–496.
- Ahumada, A., and Lovell, J. (1971). "Stimulus features in signal detection," *J. Acoust. Soc. Am.* **49**, 1751–1756.
- Bechtold, B. (2016). "bastibe/Violinplot-Matlab," Github project, available at <https://github.com/bastibe/Violinplot-Matlab> (Last viewed August 31, 2021).
- Bilsen, F. A., and Raatgever, J. (1973). "Spectral dominance in binaural lateralization," *Acoustica* **28**, 131–132.
- Blauert, J., and Allen, J. S. (1997). *Spatial Hearing: The Psychophysics of Human Sound Localization*, revised edition (MIT Press, Cambridge, MA).
- Brughera, A., Dunai, L., and Hartmann, W. M. (2013). "Human interaural time difference thresholds for sine tones: The high-frequency limit," *J. Acoust. Soc. Am.* **133**, 2839–2855.
- Brungart, D. S., Durlach, N. I., and Rabinowitz, W. M. (1999). "Auditory localization of nearby sources. II. Localization of a broadband source," *J. Acoust. Soc. Am.* **106**, 1956–1968.
- Brungart, D. S., and Simpson, B. D. (2009). "Effects of bandwidth on auditory localization with a noise masker," *J. Acoust. Soc. Am.* **126**, 3199–3208.
- Folkerts, M. L., and Stecker, G. C. (2022). "Spectral weighting functions for lateralization and localization of complex," *J. Acoust. Soc. Am.* **151**, 3409–3425.
- Good, M., Gilkey, R. H., and Ball, J. M. (1997). "The relation between detection in noise and localization in noise in the free field," in *Binaural and Spatial Hearing in Real and Virtual Environments*, edited by R. Gilkey and T. R. Anderson (Erlbaum Associates, Inc., Hillsdale, NJ), pp. 349–376.
- Good, M. D., and Gilkey, R. H. (1996). "Sound localization in noise: The effect of signal-to-noise ratio," *J. Acoust. Soc. Am.* **99**, 1108–1117.
- Henning, G. B. (1974). "Detectability of interaural delay in high-frequency complex waveforms," *J. Acoust. Soc. Am.* **55**, 84–90.
- Hintze, J. L., and Nelson, R. D. (1998). "Violin plots: A box plot-density trace synergism," *Am. Stat.* **52**, 181–184.
- Jacobsen, T. (1976). "Localization in noise," Technical Report No. 10 (Technical University of Denmark Acoustics Laboratory, Kongens Lyngby, Denmark).
- Lorenzi, C., Gatehouse, S., and Lever, C. (1999). "Sound localization in noise in normal-hearing listeners," *J. Acoust. Soc. Am.* **105**, 1810–1820.
- Macpherson, E. A., and Middlebrooks, J. C. (2002). "Listener weighting of cues for lateral angle: The duplex theory of sound localization revisited," *J. Acoust. Soc. Am.* **111**, 2219–2236.
- Moore, B. C. J., and Glasberg, B. R. (1983). "Suggested formulae for calculating auditory-filter bandwidths and excitation patterns," *J. Acoust. Soc. Am.* **74**, 750–753.
- Simpson, B. D., Gilkey, R. H., Brungart, D. S., Iyer, N., and Hamil, J. T. (2011). "The impact of masker fringe and masker spatial uncertainty on sound localization," in *Principles and Applications of Spatial Hearing*, edited by Y. Yiti Suzuki, D. Brungart, Y. Iwaya, K. Iida, D. Cabrera, and H. Kato (World Scientific, Singapore).
- Stecker, G. C., Bernstein, L. R., and Brown, A. D. (2021). "Binaural hearing with temporally complex signals," in *Binaural Hearing*, edited by R. Y. Litovsky, M. J. Goupell, R. R. Fay, and A. N. Popper (Springer International, Cham, Switzerland).
- Stecker, G. C., Ostreicher, J. D., and Brown, A. D. (2013). "Temporal weighting functions for interaural time and level differences. III. Temporal weighting for lateral position judgments," *J. Acoust. Soc. Am.* **134**, 1242–1252.
- Strutt, J. W. (1907). "XII. On our perception of sound direction," *London, Edinburgh Dublin Philos. Mag. J. Sci.* **13**, 214–232.
- Tollin, D. J., and Henning, G. B. (1999). "Some aspects of the lateralization of echoed sound in man. II. The role of the stimulus spectrum," *J. Acoust. Soc. Am.* **105**, 838–849.
- Wightman, F. L., and Kistler, D. J. (1992). "The dominant role of low-frequency interaural time differences in sound localization," *J. Acoust. Soc. Am.* **91**, 1648–1661.
- Yost, W. A. (1975). "Comments on 'Lateralization and the binaural masking-level difference' [G. B. Henning, *J. Acoust. Soc. Am.* **55**, 1259–1263 (1974)]," *J. Acoust. Soc. Am.* **57**, 1214–1215.
- Yost, W. A. (1981). "Lateral position of sinusoids presented with interaural intensive and temporal differences," *J. Acoust. Soc. Am.* **70**, 397–409.
- Yost, W. A., Wightman, F. L., and Green, D. M. (1971). "Lateralization of filtered clicks," *J. Acoust. Soc. Am.* **50**, 1526–1531.