Binaural image position distributions for phase-shifted low frequency tone bursts

Eli Osman^{a)} and Huan-yuan Tzuo

Department of Psychology, Brooklyn College, CUNY, Brooklyn, New York 11210

(Received 8 January 2010; revised 9 May 2011; accepted 12 May 2011)

This experiment was designed to yield precise measures of the statistical properties of perceived sound images. Results are reported for listeners' judgments of intracranial sound image lateral positions in response to binaural tone burst stimuli (250 Hz, 50 ms) with varying interaural phase differences, conditional on the absence or presence of a (left or right) reference monaural tone burst (also 250 Hz, 50 ms) ending 500 ms prior to the test signal. The monaural-reference shifted the position distributions toward the opposite side of the head. The position distribution variance and skewness depended on the mean of the position distribution, not on the interaural phase difference of the stimulus. The standard deviation increased as the mean moved laterally from midline. Near the midline the position distributions were skewed ipsilaterally. Near either ear they were skewed toward the midline. The results suggest that the most important noise limiting performance originates central to brainstem coincidence detector networks. © 2011 Acoustical Society of America.

[DOI: 10.1121/1.3596477]

PACS number(s): 43.66.Pn, 43.66.Ba, 43.66.Rq, 43.66.Qp [MAA]

Pages: 302–311

I. INTRODUCTION

Many binaural models have been concerned with how lateral positions of subjective images depend on interaural differences of the stimuli. Such models (see [Colburn, 1995](#)) have frequently used some form of cross-correlation of the inputs to the two ears to generate predictions of performance. One such model, the position-variable model described by [Stern and Colburn \(1978\)](#), generated a statistical variable assumed to be monotonically related to subjective lateral image position. This position is symbolized in this paper as “P,” the intracranial position represented in degrees corresponding to the azimuth angle. P was a unidimensional position variable specified statistically as having a Gaussian distribution. It was assumed to represent all of the information necessary to predict performance in binaural lateralization, interaural discrimination, and binaural detection tasks. This paper reports direct measurements of the statistical properties of P.

Many measurements of image positions of stimuli have been made. They include studies of high frequencies and variations in properties of stimulus waveforms ([Bernstein and Trahiotis, 2009, 2010](#)), interactions between interaural time delays (ITDs) and interaural level differences (ILDs) ([Domnitz and Colburn, 1977](#)), onset versus ongoing interaural disparities ([Freyman et al., 2010](#)), auditory saltation ([Getzmann, 2009](#)), and aftereffects of adaptation ([Phillips and Hall, 2005](#); [Carlile et al., 2001](#); [Vigneault-MacLean et al., 2007](#); [Haft, 1997](#)). The study reported in this paper was somewhat similar to [Yost \(1981\)](#), which extended [Sayers \(1964\)](#) and [Yost et al. \(1975\)](#). There, the listener was presented an eight unit train of tone bursts. The first three stimuli in the train had no interaural differences and served

only as a center or reference condition, whereas each of the next five stimuli had the same interaural phase delay (IPD) and served as test signals. The listeners were presented with repetitions of each eight-stimulus segment and then positioned a pointer attached to a linear slide potentiometer laid over a drawing of a head to represent image position. The potentiometer was divided into 21 equal sections to allow for the determination of 21 intracranial positions, presumably spanning 180°. This forced rounding of each judgment by $\pm 4.5^\circ$, although listeners believed the pointer scale to be continuous. During one block of trials, tonal signal frequency, overall level, duration, and type of interaural difference were all fixed. Several values of the particular type of interaural difference were presented within a block of trials, spaced at least 30° or 3 dB apart. It was found that the lateral position of the image of the test stimulus was linearly related to test stimulus IPD over the range from -90° to $+90^\circ$. Outside that IPD range images often appeared at more than one position, usually on opposite sides of the head. These results were generally independent of frequency, duration, and overall intensity. More recently, [Zhang and Hartmann \(2006\)](#) had listeners use a numerical scale to estimate image lateral positions for 1.6-s sine tones of different frequencies with different interaural delays, identified as IPDs or ITDs. The stimuli were presented in random order within 25 trial runs, and the listener could repeat a tone as many times as desired for each trial. The results showed that their listeners lateralized the sine tones on the basis of ITDs and not IPDs.¹

The intention of the present study was to measure image position variance as directly and reliably as possible. It was hypothesized that experimental conditions designed to optimize listener attention to narrowly localized subjective images of the test stimuli and minimize variance not due to those images would generally yield more precise measures of the statistical properties of P. Therefore, differing from other studies, only one brief binaural test tone of low frequency

^{a)}Author to whom correspondence should be addressed. Electronic mail: eosman@comcast.net

was presented per trial, IPDs differing only in small steps were randomized among trials, and the dependent variable P was resolved to an accuracy of approximately $\pm 0.1^\circ$. Data were collected for three experimental conditions: without a reference stimulus, with a left ear monaural-reference stimulus, and with a right ear monaural-reference stimulus. The monaural-reference stimuli were originally intended to help anchor the subjective position scale, but it was found that there was a shift in the entire set of position judgments away from the ear to which the monaural reference was presented. Such alterations of perception due to sensory adaptation have been studied extensively before. The present finding is similar to that reported in [Kashino and Nishida \(1998\)](#), where their “aftereffect” was a shift in the apparent location of the image of a test stimulus away from that of a preceding adaptor sound, although the temporal parameters of their study were quite different, and they did not use a monaural adaptor. This aftereffect is one version of “binaural adaptation,” which is the process responsible for the loss of effectiveness of binaural information of successive parts of a stimulus ([Haftor, 1997](#)). Other reports of similar effects appear in [Carlile et al. \(2001\)](#), [Phillips and Hall \(2005\)](#), and [Vigneault-MacLean et al. \(2007\)](#).

In this study both the variance and the skew of the distribution of position judgments were found to be related to mean position, rather than to the interaural difference of the stimulus. The literature contains no previous reports of similar findings. The significance of this is considered in the discussion where a qualitative model is offered to explain the results with regard to the joint literature on modeling and analysis of the nature and level of neuronal response patterns in the binaural pathways.

II. METHODS

Listeners reported their subjective lateral image position judgments for short binaural test (probe) tone bursts with interaural phase differences. A two-interval paradigm was used in which the first interval served as a reference interval and the second interval served as a test interval. The binaural signal and the monaural reference were 250-Hz tone bursts, each having a total duration of 50 ms with 10-ms rise/fall times, and the left and right binaural test stimuli were gated simultaneously so that only the sine tone and not the envelope could have an interaural delay. The choice of one presentation of a 50-ms test signal per trial was based on pilot work, as it allowed the listener to choose a single image position with confidence. The level for all stimuli was 70 dB sound pressure level (SPL) throughout the experiment; the ILDs were fixed at zero. Sixteen nominal values of IPD (θ) were used, eight leading in the left (−) ear and eight leading in the right (+) ear, with the following absolute values: 5° , 10° , 15° , 20° , 25° , 45° , 65° , and 105° . For any given IPD, on each trial, the listener’s task was to judge the lateral position of the perceived binaural image. For a particular nominal IPD, the actual IPD used was varied slightly, by $\pm 1^\circ$ or $\pm 3^\circ$, as a spacing of 2° is a rough estimate of the minimal just noticeable difference (JND) for IPD ([Mills, 1960](#); [Hershkowitz and Durlach, 1969](#); [Domnitz, 1973](#); [Yost, 1974](#)). Those four

binaural tonal signals—termed θ_i —were each presented for 25 trials within a randomly ordered 100-trial block. This was to ensure attention to each stimulus and prevent position judgments based on some anticipated value. The distribution of position judgments was determined for each of the four signals over four such blocks (400 trials). All stimulus presentations were electronically generated, routinely calibrated, and listened to independently by each of the two subjects via a matched pair of TDH-39 earphones inside a double-walled sound-proof chamber.

Two reference sounds were used: (1) a monaural tone at the left ear (i.e., whose image was fully lateralized at the left side) and (2) a monaural tone at the right ear (image fully lateralized at the right side). A third condition did not use a reference. For each of the 48 main stimulus conditions, 16 θ ’s $\times$ 3 reference conditions, the position distribution was determined from the four distributions of the $\pm 1^\circ$ and $\pm 3^\circ$ separate conditions for each listener.

The experiment was self-paced. The listener initiated each trial by closing a switch, after which there was a 1200-ms waiting period followed by a 50-ms reference interval, then a 500-ms waiting period followed by a 50-ms signal interval, and then an open response interval. The listener’s task was to position a rotatable pointer, which had a continuous semicircular scale representing positions along the surface of the head in the mid-frontal plane, so that its tip coincided with the lateral position of the subjective image projected vertically to the surface of the head between the ears (similar to [Yost, 1981](#)), press a button to record the response, return the pointer to its central location (premarked by a vertical straight line), and then push another button to reset the equipment for the next trial.

One adult female (F) and one adult male (M) served as listeners. Both had normal hearing and considerable experience in psychoacoustics tasks over the previous several years, and each had about two months intensive training for this particular experiment before the data were formally collected. Each listener was seated comfortably inside the chamber with the pointer directly in front of her/him at an angle of about 30° so that the shadow of the listener’s head formed by the light in the chamber nearly coincided with the semicircular path of the tip of the pointer. Although the second author, who was one of the listeners, ran the routines of data collection and initial processing, the first author alone did the analysis and interpretation of the data as discussed in this paper. During the data collection process neither listener could have been aware of the nature of the results regarding the relationships of variance and of skew to mean position.

The first stage of the analysis was to calculate the frequency distribution of position judgments for each value of θ_i within each combination of stimulus IPD (θ), reference condition, and listener. The means and standard deviations of these are termed $P_{Mn}(\theta_i)$ and $P_{SD}(\theta_i)$, respectively. The frequency distributions were then combined into one distribution by shifting each one laterally by the difference between the overall mean and the local mean; this effectively averages each set of moments about the mean for each set of the four choices of θ_i . The mean (P_{Mn}) and the standard deviation (P_{SD}) and the skew (G) of each listener’s

lateral position judgments of the combined distribution for each value of stimulus IPD and reference condition were then calculated.²

III. RESULTS

The results for the two listeners are shown in Fig. 1. For the no-reference condition (filled circles), the mean subjective lateral image position was monotonically related to stimulus IPD, θ . When θ was $\pm 105^\circ$, the value of P_{Mn} was in the range of 40° – 80° from subjective center, and for θ near 0° , the value of P_{Mn} approached the midline. As would be expected for low-frequency pure tones, the side of the image position was always directed toward the ear leading in phase. For the monaural-reference conditions (represented by diamonds and triangles), the monaural reference caused a shift in mean image position for the probe. This shift was away from the ear that the monaural reference was presented to: only 8 of the 64 P_{Mn} values were located on the monaural-reference's side of the midline. An ipsilateral reference shifted P_{Mn} only to just across the midline, but not much further.

The reliability of the major features of the results is supported by the correlation coefficients for the sets of four phase shifts (θ_i) surrounding each nominal value of θ and

FIG. 1. Mean lateral image position as a function of stimulus interaural phase difference. Values are negative for left ear leading in phase or position left of center, and positive for right ear leading in phase or position right of center. Data points in each panel are for the individual listeners (F, M) as indicated, for the no-reference (NR), left-reference (LR), and right-reference (RR) conditions.

their corresponding mean values for image position, $r(\theta_i, P_{Mn}(\theta_i))$. Of these 96 correlation values (16 θ 's, 3 reference conditions, 2 listeners), only five values were negative. Most of the values were close to +1; 66 values were greater than +0.8. Therefore, it was concluded that the listeners were able to judge the lateral positions of the images resulting from four binaural stimuli closely spaced by IPDs, to yield P values monotonically related to those IPDs.

Figure 2 shows the standard deviation of responses, P_{SD} , as a function of the target IPD, θ . There is no identifiable pattern for the complete set of data points. Figure 3 shows the same values³ of P_{SD} as a function of the location of the image, P_{Mn} . For $-60^\circ < P_{Mn} < +60^\circ$, P_{SD} was minimal in the region of 2° – 4° when the image was located near subjective center. P_{SD} was larger for images located laterally away from center, and essentially largest in the 10° – 15° range for images located approximately in the region beyond 10° but within 60° off center. Finally, it dropped again beyond that region, and reached 2° – 5° when the image was located at about 85° to either side of center.

Regardless of the way position judgments were shifted by the monaural reference, P_{SD} maintained its direct relation to P_{Mn} , and was not directly dependent on the reference condition. These data were fit for each listener with a segment of a sine function, shown in Fig. 3, chosen simply for convenience after visual inspection of the graphs. The function used was $P_{SD} = 7.5\sin(1.84|P_{Mn}| + 11.7) + 3.3$, with the four

FIG. 2. Standard deviation of lateral image position as a function of the interaural phase difference of the binaural test stimulus. Data points in each panel are for the individual listeners (F, M) and for the three reference conditions (NR, LR, RR), as indicated.

FIG. 3. Standard deviation of lateral image position as a function of mean lateral image position, without regard to θ . Data points in each panel are for the individual listeners (F, M) and for the three reference conditions (NR, LR, RR), as indicated. The function $P_{SD} = 7.5 \sin(1.84|P_{Mn}| + 11.7) + 3.3$ is also shown. See the text for details.

constants chosen for a best least-squares fit to the data. The variance accounted for by this fit was 53% for listener M and 60% for listener F. In contrast, however, for Fig. 2 none of the variance was accounted for by a corresponding sine-segment fit. The comparison of these two fits indicates that the change in P_{SD} due to the shifting of P_{Mn} caused by the monaural reference was related directly to P_{Mn} , but was not related to θ , the stimulus IPD.

It would be expected that the distributions of P should be more skewed toward the contralateral ear as P_{Mn} moves toward either side, due to the lateral bounds of the interaural space. This is reflected by both the monaural-reference effects reported here and the “cue reversal” phenomena (Sayers, 1964; Sayers and Cherry, 1957; Yost, 1981). To test this, the distributional shapes of $P(\theta)$ were evaluated with the χ^2 goodness-of-fit test for normality, and by computing estimates of skewness (G). For the total of 96 distributions across both listeners, it was found that 62 distributions were significantly non-normal (53 at the 0.01 level, and the other 9 at the 0.05 level). The values of skewness, G , are plotted as a function of P_{Mn} in Fig. 4. Figure 4 shows that the P distributions were more skewed toward the ipsilateral ear relative to P_{Mn} , with positive G for positive P_{Mn} and negative G for negative P_{Mn} , when P_{Mn} was closer to subjective center. They were least skewed for $|P_{Mn}|$ in the region of 45° . For the values of $|P_{Mn}| > 60^\circ$ there was a reversal of skew, toward the center of

FIG. 4. Measure of skewness as a function of mean lateral image position, in degrees. Data points in each panel are for the individual listeners (F, M) and for the three reference conditions (NR, LR, RR), as indicated.

the head. This argument is supported by the binomial or sign test for the entire set of products GP_{Mn} , which was statistically significant with one-tail $p < 0.001$. In comparison, the same test applied for the set of products $G\theta$ was not statistically significant, with one-tail $p > 0.45$. Thus, both G and P_{SD} were each dependent on P_{Mn} , but not on stimulus IPD.

IV. DISCUSSION

In summary, the results were: (1) the mean image position, P_{Mn} , for a brief 250-Hz pure tone was strongly influenced by the introduction of a monaural reference terminating one-half second before the target signal; (2) the standard deviation of responses, P_{SD} , was dependent on P_{Mn} but not on the IPD of the stimulus; (3) P_{SD} was minimal when P_{Mn} was near center, increased to maximum when P_{Mn} was about the midregion toward either ear, and decreased to another minimum when P_{Mn} was further toward either ear; and (4) the skewness, G , of the position distribution was also dependent on P_{Mn} , but not on the IPD of the stimulus, indicating boundaries at the midline, as well as at the lateral extremes of P . The apparent dependence of P_{SD} and G on P_{Mn} is the main topic in Sec. IV D. The results are first discussed in Secs. IV A–IV C in relation to lateralization, discrimination, and the binaural aftereffect.

A. Lateralization

The best linear approximations to image position as a function of stimulus IPD are informative. Figure 5 shows that for the no-reference condition, the values of P_{Mn} were

FIG. 5. Mean image position averaged over the two listeners (F and M) as a function of interaural phase difference, for the no-reference condition. Vertical bars represent standard deviations (from averaged variances) for this study. The regression line has a slope of 0.508. Results are also shown from the average over the four listeners of Yost (1981) for 200-Hz tones, from Fig. 3 of Kashino and Nishida (1998) for both the no adaptor and zero IPD adaptor conditions for 400-Hz tones, and from Lang and Buchner (2009) for 500-Hz tones.

about equal to $\theta/2$ (filled circles), i.e., the slope was about 0.5. However, for that no-reference condition the local slopes of $P_{Mn}(\theta_i)$ —i.e., the slopes for the different sets of four θ_i —averaged over the 16 θ 's and over the two listeners produced a value of 0.808. A one-tailed t -test showed this to be significantly higher ($p < 0.0005$) than the overall slope of 0.508. In other words, the $P_{Mn}(\theta_i)$ for neighboring θ_i are farther apart than would be expected from the overall function $P_{Mn}(\theta)$. This may be explained by the hypothesis that any image is somewhat displaced laterally away from the position of the image that precedes it, even within a block. Note that only neighboring θ_i were presented within each block of 100 trials in this study. Other studies of lateralization of low frequency tones (Sayers, 1964; Yost *et al.*, 1975; Yost, 1981; Kashino and Nishida, 1998; Lang and Buchner, 2009) had all IPDs presented within blocks of trials, and produced larger slopes for $P(\theta)$, as would be expected from the foregoing hypothesis (three of these data sets are replotted in Fig. 5). This idea of lateral displacement also suggests that a *centered* reference image would repel the surrounding images of other stimuli, thus producing a steeper function of $P_{Mn}(\theta)$. Within listener results for this were reported in Kashino and Nishida (1998) and are displayed in Fig. 5 for one listener (compare open triangles and open diamonds). They show such an effect (at 400 Hz) for a center-reference stimulus (IPD = 0), used as a relatively long duration adaptor. Their data shows the effect to transform the position scale by increasing its slope approximately from 0.668 (no-reference condition) to 0.816 (center reference). Varying results for their four other listeners, which were summarized in a different format, were reported to be similar to the first. Their IPD values were all within the range of $\pm 90^\circ$ and thus similar to those of this study. This brief discussion, coupled with the monaural-reference effect shown in Fig. 1, highlights the potential influence of experimental paradigm con-

text on the obtained data. Another context effect worth noting here is the one recently reported in Lang and Buchner (2009) for 1000-ms, 500-Hz tonal stimuli with a set of IPD values within the range $\pm 108^\circ$. Their averaged lateralization results for 18 listeners had a slope for $P_{Mn}(\theta)$ of 0.583. They showed that the lateral position of a sound image judged on an azimuthal scale, for which differences in one cue, IPD or ILD, were compensated by differences in the other cue, ILD or IPD, such that the sound image was judged to be at the midline position, shifted back during a lateralization task toward the location from which the sound appeared to originate before the compensation adjustment. The explanation was that attention shifted toward the effect of the to-be-adjusted cue during the compensation task, leading to a temporary increased weighting of the to-be-adjusted cue during that task. Generally, the mapping of stimulus to lateralization scale is significantly affected by experimental context within each trial and also across trials. Such effects on the slopes of lateralization functions will also have consequences for measurements of IPD JNDs depending on experimental design.

B. Discrimination

Models of binaural processing are often evaluated by comparing predictions for JNDs of interaural delay to human performance, where the variances used to generate the predictions must be inflated in order to fit the psychophysical data (e.g., Hancock and Delgutte, 2004). The authors are unaware of any study where direct measures of position variance were used to predict JNDs of interaural delay. The results by Yost (1974) for 300-ms tones showed JND functions for IPDs to be very similar for 250, 500, and 900 Hz. Those results also showed the critical variable below about 1000 Hz to be IPD rather than ITD, in agreement with Sayers (1964), although not with Zhang and Hartmann (2006). For the present experiment, JNDs were derived by assuming that the discrimination decision variable underlying the JND was the change in P (ΔP) and that the variance of the distributions was essentially constant across small shifts of the mean of P . The results reported in the literature that were for stimulus parameters most similar to those of the present study were for 300-ms tones in Yost (1974). Therefore the data for the no-reference condition (50-ms tones and $|\theta| \leq 105^\circ$) were chosen for comparison, and predictions for 300-ms tones were assumed to be related to these by the optimal square-root dependence on duration (Houtgast and Plomp, 1968); i.e., $\Delta P \cong (P_{SD})/\sqrt{6}$. The JND predictions are given by $\Delta\theta = \Delta P/m_\theta$, where m_θ is the local slope across each set of four θ_i . These values, and the JND values reported by Yost (1974) at 250 and 500 Hz, averaged over the directions left and right, are plotted as a function of interaural phase difference ($|\theta|$) in Fig. 6. The predicted and obtained JNDs are compatible.⁴

C. Aftereffect and adaptation

Kashino and Nishida (1998) studied the auditory lateralization aftereffect for low frequency tones. They varied IPDs for adaptor and for probe stimuli, in an extension of Thurlow

FIG. 6. Comparison of interaural phase difference JNDs between the predicted values from this study and the values of Yost (1974).

and Jack (1973) and other reports. Generally this work showed a shift in the apparent location of the test stimulus away from the apparent location of the adapter. They did not use a monaural adapter. The monaural reference used in the experiment reported in this paper may be considered a brief one-interval adapter, and the monaural-reference effect (Fig. 7) is similar to their aftereffect. In each of their adaptation sessions, the adapter tone was first presented for 60 s and was followed by one of the randomly selected 0.25-s test tones, after which each subsequent test tone offset was followed 4 s later by the same adapter tone for 5 s in order to reinforce the adaptation. They found the magnitude of the aftereffect to be both frequency and ITD selective, with the largest displacements (up to 20°) occurring for adapter and probe tones of identical frequency and for an ITD disparity of around $250\ \mu\text{s}$ (i.e., 36° at 400 Hz). Their intention was to find out if the selectivity of the aftereffect was consistent with the existence of ITD selective units in humans, and whether binaural models should incorporate mechanisms of

FIG. 7. Magnitude of the aftereffect for the monaural-reference (LR and RR) data of this study (averaged over the two listeners) compared to the data of Kashino and Nishida (1998) for adapter IPDs of -54° and $+54^\circ$.

adaptation. They offered a model for the aftereffect, based on coincidence detectors, using the location of peaks of a summary cross correlation function or centroid location along the ITD axis with inclusion of a gain adjustment (with no physiological interpretation) that is responsible for the shift in perceived lateral position of the binaural image. They concluded that the adaptation occurs before across-frequency integration but after binaural interaction. Studies of inferior colliculus neuron responses (McAlpine *et al.*, 2000; Ingham and McAlpine, 2004; Nelson *et al.*, 2009) have led to similar interpretations that processes central to the monaural pathways and subsequent to the level of primary binaural interaction are responsible for adaptation and forward masking. Phillips and Hall (2005) studied psychometric functions for both ITD and ILD, and found that shifts in the perceived laterality of target tones could be induced by a sequence of adaptor tones in opposite directions at two frequencies concurrently. They concluded that the effects must be mediated at or rostral to the superior olive because that is one of the first sites of binaural interaction, but prior to the decision level of perceptual judgment. They found it striking that only a few seconds of intermittent exposure to the adaptor tones was needed to produce their effects. Therefore, it is interesting that the monaural-reference effect found in the present experiment was induced by only one brief presentation of the monaural-reference stimulus. Perhaps the monaural-reference effect and the aftereffect are due to similar changes of neural activity, despite the temporal parameter differences between the psychophysical studies.

Inspired by Kashino and Nishida (1998), Furukawa *et al.* (2005) examined how single neurons in the inferior colliculus of anesthetized gerbils responded in a parallel experimental paradigm. They analyzed how a conditioner affected a single unit's responses to a probe and its IPD tuning. The dichotic stimulus sequence was a 200-ms tonal conditioner followed by a silent gap and then a 50-ms tonal probe. The conditioner and probe shared both frequency (typically the unit's best) and level (10–20 dB relative to threshold). The IPDs of the two tones were varied independently. They found that the effect of the conditioner on single unit responses to the probe vanished for gaps on the order of 400 ms. They also found that individual neurons' preferred IPDs were relatively invariant with the conditioner IPD. Using a population-vector model or a hemispheric channel model, they were able to predict that a conditioner would produce a shift in the IPD representation of the probe in a direction away from the conditioner IPD, so enhancing the contrast of the aggregate responses. They expressed concern with regard to the type and timescale of the stimuli that comparable psychophysical data showing the shift of perceived lateral position did not exist. The monaural-reference effect of this paper provides results relevant to this concern because of the 50-ms tonal stimuli employed.

D. Modeling

A qualitative model is presented here to describe the results of the experiment reported in this paper. It is represented schematically in Fig. 8. Begin with a set of

FIG. 8. (Color online) These two illustrations are schematic representations of the model described in the text. They represent binaural processing through the left brainstem only, as if viewed from behind the head. Mirror images of these illustrations would represent right brainstem processing and potential image positions from near midline and toward the left ear. In each illustration P_{IC} and P_{TH} values to the left and right represent potential image positions near the midline and toward the right ear, respectively. A right ear leading dichotic stimulus generates paired inputs to the left IC, identified by either S_{10} or S_{65} . This leads to increased firing rates for a set of neurons with varied IPD tuning curves, which are represented by the four shown in each illustration. A neuron is identified by the location along the IPD axis of the peak of its tuning curve. In each illustration, the single tuning curve to the right represents the low density of neurons with best delays in the region of 65° . The three tuning curves to the left jointly represent the higher density of neurons with best delays near 10° . The vertical long-dashed lines to the left indicate a best delay at zero IPD. Each short-dashed vertical bar represents the trial by trial standard deviation of firing rate for the neuron with best delay in the region of 65° . Each solid vertical bar represents the trial by trial standard deviation of firing rate for one of the neurons with best delay in the region of 10° . The distributions of P_{IC} over trials for each of the two stimuli feed to the thalamus as shown, where the variance of P_{IC} is larger for S_{65} and smaller for S_{10} . Neural processing through the thalamus is similar to that of the inferior colliculus, so that additional noise causes the distributions of P_{TH} to be wider than those of P_{IC} . With no-reference, NR, the increase in variance of P_{TH} is greater for S_{65} as compared to S_{10} . The effects of the monaural references are represented by the longer large arrows. The direction of each arrow symbolizes the shift of mean potential image position to result from the next level of processing. This is due to the respective temporary alterations of IPD tuning curve response strengths caused by a left- or right-reference, LR or RR, as indicated.

coincidence detector neurons in the inferior colliculus for a fixed low stimulus frequency, with their various overlapping IPD tuning curves distributed laterally across the IPD axis [bottom of Fig. 8(A)]. Each cell is identified by its best IPD, i.e., the location of the peak of its tuning curve. Let S_θ represent a binaural stimulus with an IPD value of θ . S_θ may activate each neuron with a tuning curve whose average response strength (firing rate) is positive for that IPD. On average, presentation of S_θ will generate the strongest responses for neurons with best delays at or near $IPD = \theta$, with weaker average responses for neurons with best delays displaced from θ . The actual response strength for each particular neuron will vary randomly from trial to trial (symbolized by vertical standard deviation bars superimposed over tuning curves in Fig. 8).

The maximally responding neuron will vary from trial to trial because of the variances of response strength for these neurons. The neural representation of this process on each trial is a “tentative” position variable symbolized as P_{IC} and will be the input to the next level of binaural analysis in the thalamus [midlevel of Fig. 8(A)]. The neural process through the thalamus will be similar in nature, and consequently generate an output on each trial, symbolized by P_{TH} [top of Fig. 8(A)]. A repetition of this process may occur at higher (cortical) levels of processing, but for simplicity the argument to follow will focus on the distributions of P_{IC} and P_{TH} .

To determine the mean, variance, and skew of P_{IC} , one needs to know: (a) the distribution of IC coincidence neurons identified by their best delays across IPD, (b) the distribution

of average response strength (firing rate) across stimulus IPD for each neuron (i.e., tuning curve), and (c) the distribution, or at least the variance, of firing rate for each stimulus IPD for each neuron. For simplicity, it is reasonable to assume that: (a) when arranged according to best delay the density of neurons along the IPD axis is greatest at or near 0° and decreases monotonically as IPD deviates from 0° , (b) this distribution is strongly biased for contralateral representation, so that the dominant neural response pattern is seen at only one of the two ICs, and (c) the slopes of individual IPD tuning functions at or near 0° are steeper on the side toward 0° than on the other side. These assumptions are in accord with physiological data from the mammalian auditory brain at the levels of the inferior colliculus and thalamus (Fitzpatrick *et al.*, 1997; McAlpine *et al.*, 2001). Two examples are given below to illustrate this argument. The stimulus IPD values of 10° and 65° are chosen here to exemplify features of interest, as described in the following.

Let position be based on the best delay of the maximally responsive neuron on any given trial. Across different trials the presentation of S_{10} will yield the strongest responses for one of the left IC neurons with best delays at or near 10° most of the time. The image position, P_{IC} , would accordingly be identified as the interaural correlate of that best delay, and suggest an image position near the midline. This is schematically represented by the distribution labeled P_{IC} , at the mid-level of Fig. 8(A). Sometimes however, some neuron with best delay displaced from 10° will show the strongest response due to the response strength variances of the individual neurons. Then P_{IC} would be identified as the correlate of that displaced best delay. Because of this process the trial by trial presentation of S_{10} will produce a particular distribution of image position judgments centered near the value correlated with a best delay of 10° . The variance of this distribution will be relatively small because of the higher density of neurons having best delays near 0° . Further, the P_{IC} distribution would be skewed toward larger values of best delay because of the strong bias for contralateral representation; i.e., almost all of these neurons have positive best delays and very few have negative best delays. (In Fig. 8, the relatively empty region of negative best delay is not represented.) In comparison, presentations of S_{65} will yield the strongest responses for individual neurons in the left IC with best delays at or near IPD = 65° most of the time [see Fig. 8(B)]. Then for each such trial, P_{IC} would be identified as the correlate of the best delay for the maximally responsive cell. On some trials a neuron with best delay displaced from 65° in one direction or the other will show the strongest response due to the response strength variances of the individual neurons. There would be many such cells with best delays displaced toward 0° , as well as other cells with best delays exceeding 65° . Consequently, there will be many instances of maximum response along with the correlated P_{IC} values for neurons with best delay displaced toward 0° , and others with best delay beyond 65° . Therefore, the variance over trials of P_{IC} values for a stimulus IPD of 65° would be relatively large. The P_{IC} distribution would be less skewed away from the midline because of the percentages of neurons with different best delays spread around 65° in both

directions. For stimulus IPDs much greater than 65° the skew would be reversed and the variance reduced, because there are no neurons with best delays beyond the lateral boundary, and the neurons with smaller best IPDs would determine the skew of the distribution.

There has been an ongoing debate in the literature about the best way to model binaural lateralization and discrimination (Fitzpatrick *et al.*, 1997; McAlpine *et al.*, 2001; Hancock and Delgutte, 2004). The preceding discussion is based on the “lower envelope” hypothesis, which considers the maximally responsive neuron in the collective of coincidence detectors to signal image position. In contrast, the “broad channel” hypothesis considers pooling of neural firing outputs of those detectors and comparison across hemispheres to do the same. These are essentially alternate proposals for the neural correlate of P . The model argument for the broad channel hypothesis would be similar to that given previously, except that the centroid of the neural firing rate (response strength) versus IPD function would be the correlate of P_{IC} . If sufficient neural data were available to make more detailed quantitative assumptions, it is possible that the statistical properties of the psychophysical results could favor one of these hypotheses over the other.

It may be that if the lateral shift of P caused by the monaural reference were explained by some gain adjustment function of IPD, temporarily affecting the relative response strengths of unit IPD tuning curves, as in Kashino and Nishida (1998) or Furukawa *et al.* (2005), then predictions for variance and skewness of P_{IC} might match the results of Figs. 3 and 4. However, with regard to the 500-ms reference-test gap of the study reported in this paper, the results of other studies (Zwislocki, 1978; Brosch and Schreiner, 1997; Meddis and O’Mard, 2005; Harris and Dallos, 1979) render it very unlikely that the monaural-reference effect can be explained by a peripheral neural process. In particular, the single neuron results of Furukawa *et al.* (2005) imply that it must be mediated central to the IC, beyond the level of those coincidence detectors. Even for single units in the primary auditory cortex (anesthetized cats) the longest post masker period yielding reduction of responses to the test tone burst of either a 30- or 50-ms duration was in the range of 53–430 ms (Brosch and Schreiner, 1997).

The IC feeds to the next level of synaptic processing, the thalamus, where the variability of the subsequent distribution of tentative image position values, “ P_{TH} ” or output from the thalamus for this example, is determined. The upper level of each of Figs. 8(A) and 8(B) illustrates this portion of the argument. Suppose that at this level the monaural reference produces the gain adjustment as described previously [which is represented by the longer large arrows in Figs. 8(A) and 8(B)]. Here it is assumed that the gain adjustment is a consequence of temporarily altered patterns of synaptic excitation and inhibition and thus effectively altered IPD tuning curves of thalamic neurons. This essentially reduces response strength for neurons with best delays favoring the side of the monaural reference. Over those trials where P , as psychophysically determined, is shifted laterally due to the monaural reference, the variability of P_{IC} would nevertheless still be in accord with the P_{IC} variability already determined.

Additional neural noise would be generated at the next level of processing to yield the distribution of P_{TH} values. The following is based on estimates from visual inspection of the psychophysical trends of Figs. 1 and 3. A stimulus IPD that produces P values around 5° azimuth without a monaural reference produces P values around 30° azimuth with a left side monaural reference, and they have a much higher variance over trials. Moreover, a stimulus IPD that produces P values around 30° azimuth without a monaural reference produces P values near 0° azimuth with a right side monaural reference, and they have a much lower variance over trials. One possible explanation for this is the following. The higher (thalamic) level of neural processing, which may further sharpen the neural representation of image position, has a distribution of neural best delays along the IPD axis that is analogous to that of the IC, and, after the gain adjustment across best delays, the intrinsic noise generated at that level is the primary determinant of the resulting variance of the position distribution. The distribution shape for P_{TH} as reflected by skewness may be comparably determined.

Following the arguments given previously, the relative contributions of neural noise to perceived position, as represented in outputs of the IC network and the subsequent outputs of higher level binaural networks, should show the latter to be dominant. In the Stern and Colburn (1978) position-variable model, the only source of variability derived from the random character of auditory nerve firing patterns. Recently, Kripi and Furst (2009) simulated the firing patterns of excitatory–excitatory and excitatory–inhibitory coincidence detector cells to determine their capabilities to discriminate interaural phase delays (IPDs) and interaural level differences (ILDs) for pure tone binaural stimuli. Their results were identified as “lower bounds” on both performance and model estimator variances. Fitzpatrick *et al.* (1997) showed further sharpening of neuronal IPD tuning to occur in the thalamus, where distributions of average response strength versus best delay are considerably narrower as compared to those of the IC. Their simulations for JNDs of ITD, based on sets of up to 40 randomly selected neurons, showed improvement from SOC (superior olivary complex) to IC to thalamus. When describing a neural pooling model of ITD discrimination based on the response properties of ITD-sensitive cells in the IC (anesthetized cats), Hancock and Delgutte (2004) incorporated an efficiency parameter with no direct physiological correlate, which served to effectively inflate each estimated individual neuron firing rate variance for the model by a factor of 18. As they pointed out, it is not possible to characterize all sources of inefficiency (intrinsic noise).

The qualitative argument presented here is compatible with another feature of the monaural-reference effect as it appears in Fig. 1. Relative to stimulus IPD and P values with no reference stimulus, image positions were shifted even further toward the near ear by a contralateral reference. This follows from the model because of the alteration of the tuning curves in the contralateral thalamic nucleus as explained earlier. The ipsilateral thalamic nuclei would have little representation for these stimuli. Further, image positions for larger IPDs were shifted toward the midline, but not too much further, by an ipsilateral reference, yielding the flatten-

ing shown in Fig. 1. This follows from the heavy contralateral representation and minimal ipsilateral representation by the neural tuning curves, with the understanding that any monaural reference has its potential influence exercised bilaterally.

V. CONCLUSION

This paper reports an experiment designed to measure the statistical properties of lateral image position judgments for low frequency binaural tone bursts. The stimulus parameters and procedures were chosen to best allow listener responses to reflect the image position for each individual trial, and to minimize the variance not attributable to properties of the images themselves. It was found that a monaural-reference sound shifted the image positions of the targets contralaterally, and that the variance and the skew of the image positions depended on the mean of the position distribution, rather than on the interaural phase difference of the stimulus. These effects may result from a gain function affecting the tuning response properties of binaural coincidence-detection neurons at or beyond the midbrain level. Generally, the mapping of stimulus to lateralization scale was significantly influenced by experimental paradigm.

ACKNOWLEDGMENTS

H-y.T. prepared and ran the data collection process, and both he and Pi-fong Lin Tzuo worked diligently as the two listeners. The work and guidance provided by the reviewers of previous versions of this manuscript are truly very much appreciated.

¹Although ITDs may be the more fundamental parameter for describing human lateralization and localization in general, only 250-Hz tones with absolute values of IPD not more than 105° were presented in the experiment reported in this paper, and each stimulus value can be readily converted by the proportional relationship between IPD and ITD within one cycle of a tonal stimulus.

²The differences in variance within each set of four distributions of $P(\theta_i)$ were presumed to be relatively small. F-ratio tests for homogeneity of variance (one-tailed 0.01 type-one error rate) were nonsignificant for 60 of these 96 sets (16 θ 's, 3 reference conditions, 2 listeners). For the remaining 36, reruns of each test after omitting one extreme value showed 31 to be nonsignificant. The variance adjustment required for removal of the effects of mechanical skill and equipment problems was small, with 85%–98% of the variance in the data attributable to the subjective process of lateralization of the binaural image. As it took about 20 min for a listener to complete each block of 100 trials, listener fatigue might have affected performance. Two t-tests were run for each of the two listeners, comparing the means and the standard deviations resulting from the first 33 and the last 33 trials for each of 15 data blocks, where five blocks were selected randomly for each reference condition over all values of θ . All four were insignificant (for a two-tailed 0.10 type-one error rate). The individual distributions of P for $\theta = \pm 5^\circ$, with no reference stimulus, for both listeners, showed a sufficient number of crossings over the midline, $P = 0^\circ$, implying that the listeners did not use the center vertical line (at the top of the semicircular path) as a strict boundary for their position judgments. Additional data were collected (no-reference condition and $\theta = -65^\circ, -15^\circ, +5^\circ$, and $+45^\circ$) without that center line, but with two lines at $\pm 45^\circ$ away from center. An analysis of variance for repeated measurements confirmed that the four very small differences of PSD for this procedure and the corresponding data of the main study were insignificant.

³For the two monaural-reference conditions the 64 pairs of values of P_{Mn} and P_{SD} were examined (16 θ 's $\times$ 2 monaural-references $\times$ 2 listeners). For 19 of the 64, P_{Mn} was located within 12° of the midline. Of these, 15

- values of P_{SD} were less than their associated values (for common IPDs) for the no-reference condition. By a binomial or sign test this result is statistically significant with one-tail $p < 0.01$.
- *Regarding discrimination, the Yost (1974) results as reproduced in Fig. 6 show the JNDs to be largest for $|θ| > 105^\circ$. The results of the current study suggest that in this region P_{SD} may be relatively small with the P distribution skewed toward the midline, probably due to the boundaries at the two ears. The large JNDs may be explained by the Yost (1981) findings as reproduced in Fig. 5: that $P_{Mn}(θ)$, now outside the range of linearity, is flattened.
- Bernstein, L. R., and Trahiotis, C. (2009). "How sensitivity to ongoing interaural temporal disparities is affected by manipulations of temporal features of the envelopes of high-frequency stimuli," *J. Acoust. Soc. Am.* **125**, 3234–3242.
- Bernstein, L. R., and Trahiotis, C. (2010). "Accounting quantitatively for sensitivity to envelope-based interaural temporal disparities at high frequencies," *J. Acoust. Soc. Am.* **128**, 1224–1234.
- Brosch, M., and Schreiner, C. E. (1997). "Time course of forward masking tuning curves in cat primary auditory cortex," *J. Neurophysiol.* **77**, 923–943.
- Carlile, S., Hyams, S., and Delaney, S. (2001). "Systematic distortions of auditory space perception following prolonged exposure to broadband noise," *J. Acoust. Soc. Am.* **110**, 416–424.
- Colburn, H. S. (1995). "Computational models of binaural processing," in *Auditory Computation*, edited by H. Hawkins and T. McMullin (Springer, New York), pp. 332–400.
- Domnitz, R. H. (1973). "The interaural time jnd as a simultaneous function of interaural time and interaural amplitude," *J. Acoust. Soc. Am.* **53**, 1549–1552.
- Domnitz, R. H., and Colburn, H. S. (1977). "Lateral position and interaural discrimination," *J. Acoust. Soc. Am.* **61**, 1586–1598.
- Fitzpatrick, D. C., Batra, R., Stanford, T. R., and Kuwada, S. (1997). "A neuronal population code for sound localization," *Nature* **388**, 871–874.
- Freyman, R. L., Balakrishnan, U., and Zurek, P. M. (2010). "Lateralization of noise-burst trains based on onset and ongoing interaural delays," *J. Acoust. Soc. Am.* **128**, 320–331.
- Furukawa, S., Maki, K., Kashino, M., and Riquimaroux, H. (2005). "Dependency of the interaural phase difference sensitivities of inferior collicular neurons on a preceding tone and its implications in neural population coding," *J. Neurophysiol.* **93**, 3313–3326.
- Getzmann, S. (2009). "Exploring auditory saltation using the 'reduced-rabbit' paradigm," *J. Exp. Psychol.: Human Percept. Perf.* **35**, 289–304.
- Haft, 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 (Erlbaum, Mahwah, NJ).
- Hancock, K. E., and Delgutte, B. (2004). "A physiologically based model of interaural time difference discrimination," *J. Neurosci.* **24**, 7110–7117.
- Harris, D. M., and Dallos, P. (1979). "Forward masking of auditory nerve fiber responses," *J. Neurophysiol.* **42**, 1083–1107.
- Hershkowitz, R. M., and Durlach, N. I. (1969). "Interaural time and amplitude JNDs for a 500-hz tone," *J. Acoust. Soc. Am.* **46**, 1464–1467.
- Houtgast, T., and Plomp, R. (1968). "Lateralization threshold of a signal in noise," *J. Acoust. Soc. Am.* **44**, 807–812.
- Ingham, N. J., and McAlpine, D. (2004). "Spike-frequency adaptation in the inferior colliculus," *J. Neurophysiol.* **91**, 632–645.
- Kashino, M., and Nishida, S. (1998). "Adaptation in the processing of interaural time differences revealed by the auditory localization aftereffect," *J. Acoust. Soc. Am.* **103**, 3597–3604.
- Krips, R., and Furst, M. (2009). "Stochastic properties of auditory brainstem coincidence detectors in binaural perception," *J. Acoust. Soc. Am.* **125**, 1567–1583.
- Lang, A. G., and Buchner, A. (2009). "Relative influence of interaural time and intensity differences on lateralization is modulated by attention to one or the other cue: 500-Hz sine tones," *J. Acoust. Soc. Am.* **126**, 2536–2542.
- McAlpine, D., Jiang, D., and Palmer, A. R. (2001). "A neural code for low-frequency sound localization in mammals," *Neuroscience* **4**, 396–401.
- McAlpine, D., Jiang, D., Shackleton, T. M., and Palmer, A. R. (2000). "Responses of neurons in the inferior colliculus to dynamic interaural phase cues: evidence for a mechanism of binaural adaptation," *J. Neurophysiol.* **83**, 1356–1365.
- Meddis, R., and O'Mard, L. P. (2005). "A computer model of the auditory-nerve response to forward-masking stimuli," *J. Acoust. Soc. Am.* **117**, 3787–3798.
- Mills, A. W. (1960). "Lateralization of high-frequency tones," *J. Acoust. Soc. Am.* **32**, 132–134.
- Nelson, P. C., Smith, Z. M., and Young, E. D. (2009). "Wide-dynamic-range forward suppression in marmoset inferior colliculus neurons is generated centrally and accounts for perceptual masking," *J. Neurosci.* **29**, 2553–2562.
- Phillips, D. P., and Hall, S. E. (2005). "Psychophysical evidence for adaptation of central auditory processors for interaural differences in time and level," *Hear. Res.* **202**, 188–199.
- Sayers, B. McA. (1964). "Acoustic-image lateralization judgments with binaural tones," *J. Acoust. Soc. Am.* **36**, 923–926.
- Sayers, B. McA., and Cherry, E. C. (1957). "Mechanism of binaural fusion in the hearing of speech," *J. Acoust. Soc. Am.* **29**, 973–987.
- Stern, R. M., Jr., and Colburn, H. S. (1978). "Theory of binaural interaction based on auditory-nerve data. IV. A model for subjective lateral position," *J. Acoust. Soc. Am.* **64**, 127–140.
- Thurlow, W. R., and Jack, C. E. (1973). "Some determinants of localization-adaptation effects for successive auditory stimuli," *J. Acoust. Soc. Am.* **53**, 1573–1577.
- Vigneault-MacLean, B. K., Hall, S. E., and Phillips, D. P. (2007). "The effects of lateralized adaptors on lateral position judgements of tones within and across frequency channels," *Hear. Res.* **224**, 93–100.
- Yost, W. A. (1974). "Discrimination of interaural phase differences," *J. Acoust. Soc. Am.* **55**, 1299–1303.
- 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., Tanis, D. C., Nielsen, D. W., and Bergert, B. (1975). "Interaural time vs. interaural intensity in a lateralization paradigm," *Percept. Psychophys.* **18**, 433–440.
- Zhang P. X., and Hartmann, W. M. (2006). "Lateralization of sine tones—interaural time vs phase," *J. Acoust. Soc. Am.* **120**, 3471–3474.
- Zwislocki, J. J. (1978). "Masking: Experimental and theoretical aspects of simultaneous, forward, backward, and central masking," in *Handbook of Perception*, Vol. IV. Hearing, edited by E. C. Carterette and M. P. Friedman (Academic, New York), 283–336.