Localization dominance in the median-sagittal plane: Effect of stimulus duration

Roberto M. Dizon and Ruth Y. Litovsky[a]
Hearing Research Center, Boston University, Boston, Massachusetts 02215

(Received 19 June 2003; accepted for publication 22 March 2004)

Localization dominance is an aspect of the precedence effect (PE) in which the leading source dominates the perceived location of a simulated echo (lagging source). It is known to be robust in the horizontal/azimuthal dimension, where binaural cues dominate localization. However, little is known about localization dominance in conditions that minimize binaural cues, and most models of precedence treat the phenomena as ''belonging'' to the binaural system. Here, localization dominance in the median-sagittal plane was studied where binaural cues are greatly reduced, and monaural spectral/level cues are thought to be the primary cues used for localization. Lead–lag pairs of noise bursts were presented from locations spaced in 15° increments in the frontal, median-sagittal plane, with a 2-ms delay in their onsets, for source durations of 1, 10, 25, and 50-ms. Intermixed with these trials were single-speaker trials, in which lead and lag were summed and presented from one speaker. Listeners identified the speaker that was nearest to the perceived source location. With single-speaker stimuli, localization improves as signal duration is increased. Furthermore, evidence of elevation compression was found with a dependence on duration. With lead–lag pairs, localization dominance occurs in the median plane, and becomes more robust with increased signal duration. These results suggest that accurate localization of a co-located lead–lag pair is necessary for localization dominance to occur when the lag is spatially separated from the lead. © *2004 Acoustical Society of America.* [DOI: 10.1121/1.1738687]

PACS numbers: 43.66.Qp, 43.66.Rq, 43.66.Pn [AK] Pages: 3142–3155

I. INTRODUCTION

Sounds generated in reverberant rooms produce multiple reflections that arrive from hard surfaces such as walls, and that contribute to the spatial character of the sound. However, they have a surprisingly small effect on source localization when the source contains a well-defined onset (Hartmann, 1983). This phenomenon has commonly been referred to as the ''precedence effect'' (PE) (Wallach *et al.*, 1949; Zurek, 1980) or ''law of the first wavefront'' (Blauert, 1997; Litovsky *et al.*, 1999). The PE has gained interest since it is thought that the auditory system may perform specialized processing to achieve this performance. The PE has been described as resulting from a temporary reduction in sensitivity to localization information contained in reverberation following the onset of a source. As such, in a simple paradigm whereby a source (lead) and single echo (lag) occur, perceived location of the lead–lag pair is dominated by the localization information associated with the leading stimulus. For click stimuli, this dominance is most robust when the lagging stimulus occurs within a few ms of the lead.

The PE and related phenomena have been a topic of interest for over half a century (for review see Litovsky *et al.*, 1999), although most of what is known relates to temporal characterization of the phenomena, rather than spatial variables, including the locations of the source and simulated reflections. Furthermore, experiments have typically utilized or simulated stimuli that occur in the azimuthal plane, where changes in source locations are associated with clearly defined and often perceptible differences in binaural cues such as interaural differences in time and level. Cue manipulation in the azimuthal dimension has been preferred since interaural cues map precisely to azimuthal locations. In addition, they are easily generated, as well as replicated and presented in a realistic manner to subjects. An additional benefit of presenting sounds in the azimuthal dimension is that results can be compared with what is known about neurophysiological activity in the auditory pathway in a relatively straightforward manner (Yin, 1994; Fitzpatrick *et al.*, 1995; Litovsky *et al.*, 1997a; Litovsky and Yin, 1998a,b). Finally, models of the PE that can successfully predict performance rely on interaural cues that are available in the azimuthal dimension (e.g., Lindemann, 1986; Shinn-Cunningham *et al.*, 1993; Hartung and Trahiotis, 2001; Tollin, 1998).

In contrast, the processes involved in determining the location of sound sources occurring in the median-sagittal plane are more poorly understood and little is known about the PE in the median plane. This problem is especially interesting because localization in the median-sagittal plane is mediated primarily by spectral filtering by the pinnae, head and torso of stimuli reaching the ears from various elevations (Searle *et al.*, 1975; Gardner and Gardner, 1973; Butler, 1969; Hebrank and Wright, 1974; Middlebrooks and Green, 1991). While it has been shown that binaural cues cannot be ruled out as contributing to median plane localization (e.g., Middlebrooks, 1992; Hofman and Van Opstal, 2003), a significant amount of evidence suggests that monaural spectral cues may be primary. To date, little is known about the PE in the median plane and the extent to which directional infor-

---

[a] Present address: University of Wisconsin, 1500 Highland Avenue, Madison, WI 53705; electronic mail: Litovsky@waisman.wisc.edu.

mation in reflections can be outweighed by information contained in the source when spectral cues are the primary cue for localization.

One study reported that fusion echo thresholds measured in the median sagittal plane are similar to those found in the azimuthal plane (Rakerd *et al.*, 2002). Two studies have attempted to measure dominance of the leading source in localization for stimuli in the median-sagittal plane. In a brief report, Blauert (1971) suggested that the leading stimulus dominates localization for inter-stimulus intervals of 550 μs, but not for smaller delays (within the summing localization range). Litovsky *et al.* (1997a) found dominance by the lead location for delays up to 5-ms (within the range of the PE), with diminished dominance at longer delays. In both studies however, source locations were limited to front, back, and overhead, where localization of single source sounds is difficult to interpret, both due to front–back errors and to the fact that sounds presented overhead are very poorly localized. There are also physiological data which suggest that the strength of echo suppression in the responses of single neurons in the inferior colliculus is highly similar in the azimuthal and median planes (Litovsky *et al.*, 1997a; Litovsky and Yin, 1998a). Although that does not predict the relative strength of localization dominance in the two planes, it provides further evidence for the existence of precedence phenomena in the median plane. Our hypothesis, that well-localized sounds should produce localization dominance regardless of which directional cues are being utilized, could not be affirmed by previous work.

In the present study, we selected six locations, all in the frontal hemifield on the median-sagittal plane, where single-source noise bursts were well localized by all subjects. Lead and lag stimuli were presented from various combinations of these locations, and the effect of source duration was also explored. Using these parameters we tested the hypotheses that localization dominance is robust in the median-sagittal plane, providing that a co-located lead and lag stimulus produces a well-localized image.

II. METHODS

A. Subjects

Four subjects (two male, two female) between the ages of 18 and 24 participated in the experiments. Each had some prior experience in sound localization experiments. However, all subjects were naive as to the nature of the stimuli and the goals of the experiment. Each was given at least an hour of practice with feedback to become familiar with the experimental setup and paradigm before the start of the experiment. Subjects were also given 10 min of practice with feedback prior to each session. All subjects had normal hearing as verified by a standard audiometric threshold exam.

B. Apparatus

The experiments were performed in an anechoic chamber. The subject was seated with his or her head at the center of a circular arc approximately 2.4 m in diameter mounted vertically and positioned such that it was aligned with the subject's median-sagittal plane. Six speakers composed of a single 6.35 cm driver in a sealed enclosure $(7.9\,\text{cm}\,H \times 7.6\,\text{cm}\,W \times 12\,\text{cm}\,D)$ were mounted at 15° increments between $-30°$ and 45°. The speakers produced a flat response from 300 Hz to 15 kHz. Variations in frequency response among the speakers were minimal and not compensated for. The speakers were visible to the subjects and were labeled 1 through 6 as shown in the Fig. 1.

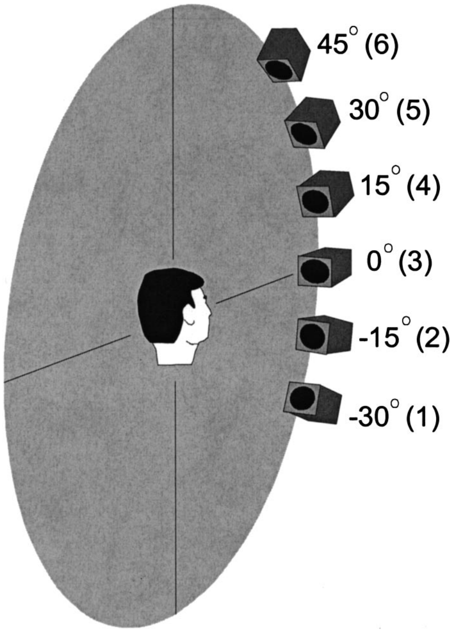

FIG. 1. Speakers are placed on a circular ring 2.4 m in diameter in the positions shown. The subject is seated with his/her head in the center of this ring.

The subject's head was constrained by a headrest mounted on the rear of the seat (Whitmeyer Biomechanix Soft-2S). The headrest has adjustable padded supports on the back of the head and under the jaws that provide support but which are not constrictive, while also being minimally acoustically obtrusive. The subjects were also told to keep their head still during stimulus presentation. They were however free to look at the speakers, which were all within the subject's visual field.

Hardware including Tucker Davis Technologies (TDT) System II hardware (AP2,DD1,PM1) in conjunction with a PC host, was responsible for stimulus computation and generation, control of the multiplexer for speaker switching, communication with the response terminal, and used as the user interface for the experimenter. The direct sound and simulated reflection signals from the D/A converter (sampling rate 50 kHz) were amplified by a Crown D-75 amplifier which was calibrated for equal gain to both channels. The amplified signals were directed to the appropriate speaker(s) with a multiplexer (TDT PM1).

C. Stimuli

Measurements were made on single-speaker and precedence (localization dominance) trials. On each trial, stimuli were comprised of a train of four identical broadband noise bursts, whose onsets were spaced 250-ms apart, as shown in Fig. 2. Noise bandwidth was effectively limited only by the

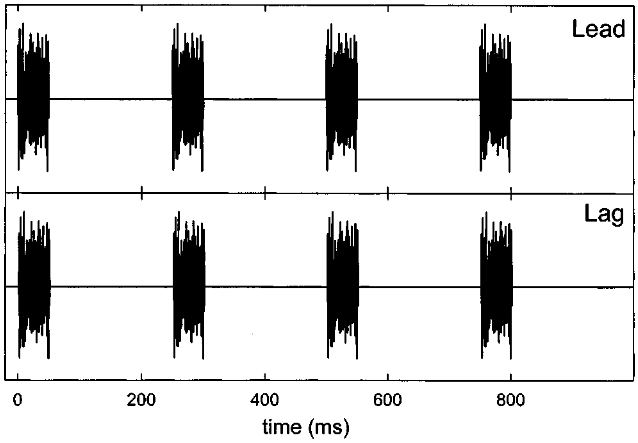

FIG. 2. Plot of sample lead and lag stimuli for the precedence conditions. This example includes 50-ms bursts, hence the 2-ms delay might be difficult to discern on this scale. All bursts within each train are identical, as are the lead and lag bursts. Lead and lag stimuli differ only in the onset delay of the lag.

speaker drivers, which are flat to 15 kHz. A different sample of noise was used for every trial. Onset and offset ramps were not applied, resulting in abrupt onsets and offsets since the broadband bursts were presented at full bandwidth by the playback system. On precedence trials, two trains were presented from separate speakers; lead and lag noise bursts were always identical to one another except for the 2-ms onset delay in the lag. Noise burst durations were 1, 10, 25, and 50-ms. Since the lead–lag delay was always 2-ms, the lead and lag bursts overlapped in time for durations of 10, 25, and 50-ms but not for the 1-ms duration. For the single-speaker trials, the lead and lag stimuli were digitally summed prior to presentation from the single speaker. The lag is included in the single-speaker trials so that differences between single-speaker and PE performance can be attributed to lag location specifically, as opposed to both the presence and the location of the lag. It is important to note that the spectral comb filtering created by the addition of the delayed repetition would occur for the single-speaker trials as well as the precedence trials.

Presentation levels were chosen for each subject individually. Detection thresholds were first established for the 1-ms stimulus presented from speaker 3 (directly in front). Subsequent presentations of 1-ms bursts were presented at 40 dB above this level, while the amplitudes of the longer duration signals were digitally attenuated by the square root of their durations to provide some degree of loudness compensation.

For the precedence paradigm trials, three pairs of speaker positions were used. These were position pairs 1 and 5, 2 and 5, and 2 and 6, using the position numbering convention shown in Fig. 1. These pairs contain angular separations of 60° (1-5 and 2-6), or 45° (2-5). These wide separations were chosen based on pilot data (Dizon *et al.*, 1997), which suggested that separations as wide as these were necessary to observe the influence of the PE, given the decrease in localization precision for the PE stimuli in that study. The three position pairs, along with their alternate order equivalents, result in six combinations of lead and lag positions. These, combined with all six single-speaker trials (lead and lag from the same speaker), result in twelve possible position types. Each position type was presented at each of the four durations 30 times, for a total of 1440 trials per subject. Trials were mixed randomly and presented in blocks of 100 (with one block of 40). Hence, single-speaker localization was measured within the same blocks as the precedence trials.

D. Testing protocol

Responses were made using a small handheld response terminal (QSI Qterm II). The subject's task was to identify a speaker that was nearest to the perceived location of the auditory image. If the stimulus appeared to emanate from more than one speaker or from a location other than one of the six possible locations, instructions were to choose the one speaker that appeared to be most salient and to contain the majority of the sound image. No feedback was provided during the experimental runs. However, subjects were trained on the single-speaker condition, and given feedback during a 1 h training session before the experiments begun, as well as for 10 min at the beginning of each session.

The forced choice protocol was chosen based on the results of an earlier median plane study (Dizon *et al.*, 1997) in which a more unconstrained response method produced a high inter-subject variability in the mapping of their perceptions onto the response choices. In that study, subjects were permitted to choose either one or two locations, depending on which one better described their percept. One subject chose two locations almost exclusively, while another chose two locations only twice out of 1500 trials. Given the variability in the data, the poor evidence of localization dominance for the nonfused judgments, and the informal comments regarding the vagueness of the stimuli, it was decided that a simple and constrained response method would best uncover a bias toward the leading stimulus. Finally, the identification paradigm was selected in an attempt to maximize any effects of localization dominance, bearing in mind the fact that identification paradigms are easier for subjects than unconstrained localization paradigms. One of the motivations for this study (see Sec. I) was to extend an earlier report by Litovsky *et al.* (1997a) in which an identification paradigm was used with only three source positions, including directly overhead.

III. SINGLE-SPEAKER RESULTS AND DISCUSSION

A. Results

Single-speaker results are shown in Fig. 3. Each row comprises data from one subject and results for the four durations are organized according to columns. These data correspond to trials in which both the lead and the lag emanated from the same speaker. Data are presented as confusion matrices, where the *x*-axis corresponds to the actual speaker number and the *y*-axis corresponds to the subject's response. The area of each closed circle is proportional to the number of responses for each condition.

Perfect source identification performance would correspond to subject responses exclusively consistent with the

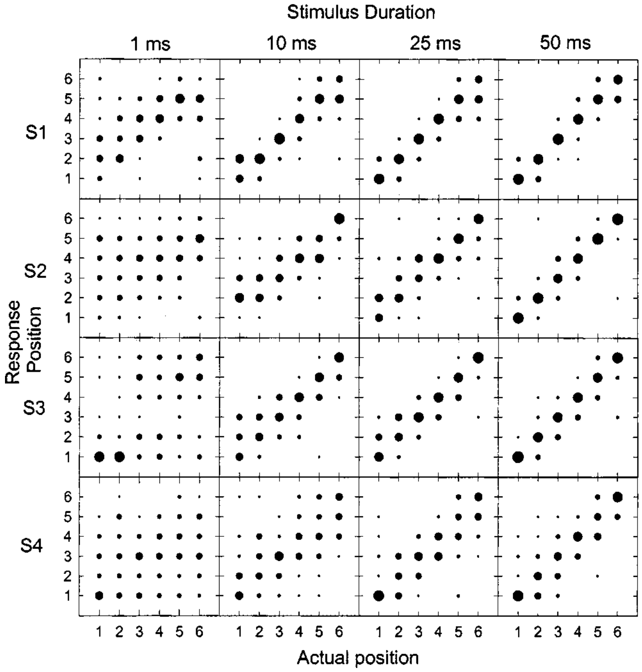

FIG. 3. Results for the single-speaker trials. Each row comprises data from one subject and results for the four durations are organized according to columns. The *x*-axis labels correspond to the actual speaker number and the *y*-axis labels correspond to the subject's response. The area of each closed circle is proportional to the number of responses for each condition.

source speaker, which would appear as maximal response frequencies along the positive diagonal in each panel. Comparing single-speaker performance across durations for each subject reveals that identification accuracy improves as burst duration is increased. Specifically, there is only a weak trend toward the diagonal in the responses for the 1-ms duration, with variation in the trend between individual subjects, while there is a strong trend toward responses near the diagonal for the 50-ms duration case for each subject.

Precision of responses improves as burst duration is increased, which can be seen by qualitatively comparing the "spread" in the responses across the speaker positions for each burst duration. The response distributions for the 50-ms case are much more tightly clustered than those in the 1-ms cases. Also noticeable is the inter-subject variability, which is most evident in the 1-ms duration data. Subjects 1 and 3 appear to be able to extract some directional information at this duration, while subjects 2 and 4 are less able to do so. These data were exposed to a number of statistical analyses to allow more quantitative descriptions of trends in the data.

The response mean $M[k]$ was computed for each subject as the mean of that subject's responses when speaker k was presented, and is given by

$$M[k]=\frac{1}{N}\sum_{n=1}^{N} r_{n,k},$$

where $r_{n,k}$=response to trial n when speaker k was presented (if the subject responded "4" when speaker 2 was presented for the 25th time, then $r_{25,2}=4$). N=the number of presentations at each speaker (30).

The response standard deviation $\sigma[k]$ for each subject is the standard deviation from the mean response $M[k]$ expressed in degrees,

$$\sigma[k]=\sqrt{\frac{A^2}{N}\sum_{n=1}^{N}(r_{n,k}-M[k])^2}.$$

where A=the speaker spacing in degrees (15°).

The rms error $E[k]$ for each subject is the standard deviation from the actual speaker number k expressed in degrees, as given by

$$E[k]=\sqrt{\frac{A^2}{N}\sum_{n=1}^{N}(r_{n,k}-k)^2}.$$

Additionally, an average of each of these statistics over the four subjects was computed, which are denoted as $\langle M[k]\rangle$, $\langle\sigma[k]\rangle$, and $\langle E[k]\rangle$.

These across-subject averages are shown in Fig. 4. In each panel, data for each burst duration are shown and differentiated from each other with symbols as indicated in the legend. Also within each panel, a sub-panel is shown which represents the standard deviation of that statistic averaged over the four subjects for each duration. A "chance" statistic is also shown within each panel using a dashed line, and is computed from a hypothetical response distribution representing chance performance, in which responses to the 30 trials for each source speaker position are distributed evenly over the six possible responses.

Considering the response mean $\langle M[k]\rangle$ in Fig. 4(a), the slopes of these mean curves start out shallow for the shortest duration, and approach a slope of 1 as duration is increased. Perfect performance would correspond to a slope of 1, while chance performance is indicated with the dashed horizontal line, and is simply a mean of the angles of the six speakers. Based on $\langle M[k]\rangle$ alone, there appears to be a compression of perceived elevation that is more pronounced as burst duration is shortened. However, without considering the other statistics, such as the standard deviation, it is difficult to make a claim of a perceived compression of elevation.

The standard deviation $\langle\sigma[k]\rangle$ is shown in Fig. 4(b), and represents the deviation from the mean $M[k]$, for individual subjects at each duration. $\langle\sigma[k]\rangle$ gets smaller as duration is increased, with the largest jump in $\langle\sigma[k]\rangle$ between 1 and 10-ms. In addition, $\langle\sigma[k]\rangle$ is of similar magnitude across the six positions for each of the durations. All values of $\langle\sigma[k]\rangle$ are less than the chance statistic, which is fixed at 27.5°. Based on $\langle\sigma[k]\rangle$, localization precision improves as burst duration is increased. This is evident qualitatively in the raw response distributions shown in Fig. 3.

The rms error $\langle E[k]\rangle$ is shown in Fig. 4(c). Similar to $\langle\sigma[k]\rangle$, $\langle E[k]\rangle$ decreases as duration is increased. However, unlike $\langle\sigma[k]\rangle$, $\langle E[k]\rangle$ increases toward either edge of the set of responses for the 1-ms case, and also somewhat for the 10-ms case, while $\langle\sigma[k]\rangle$ remained approximately constant with k relative to the response set. Since $\langle E[k]\rangle$ is referenced to the actual speaker presented, this statistic is a reflection of the compression in $\langle M[k]\rangle$, which is more pronounced for the shorter durations.

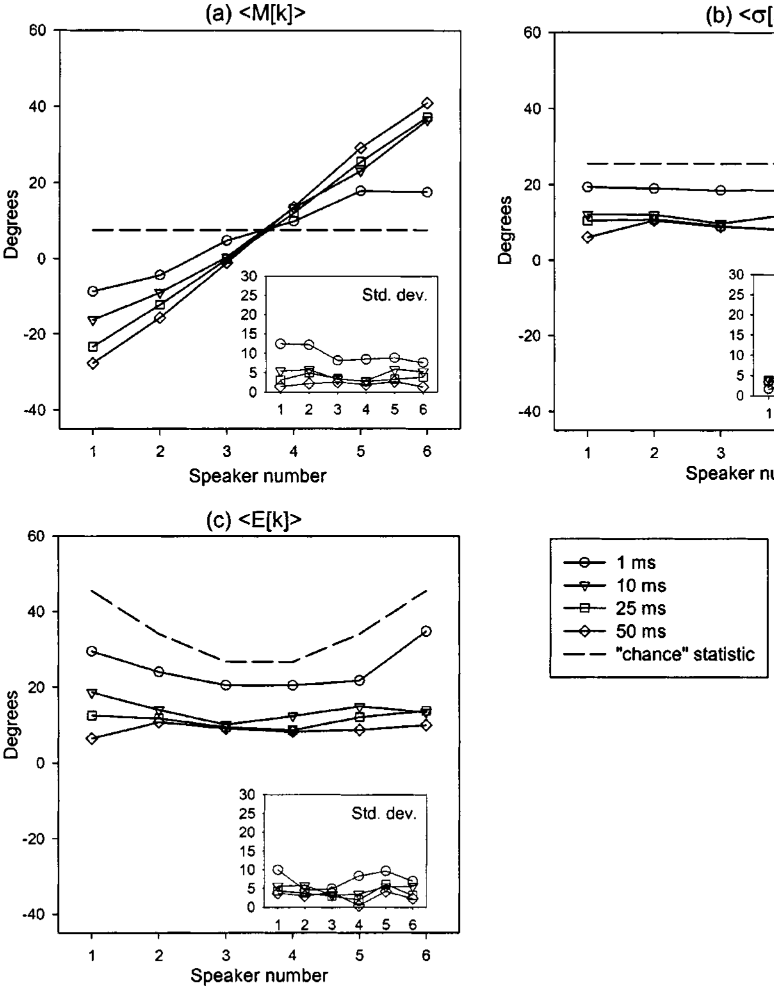

FIG. 4. (a) Mean response in degrees $M[k]$, (b) standard deviations $\langle\sigma[k]\rangle$, and (c) rms error $\langle E[k]\rangle$ for the single-speaker condition plotted against actual speaker number for each of the four durations as indicated in the legend. Each data point represents a mean of the statistic over the four subjects. The standard deviations about each mean are plotted in the inset of each panel. The dashed line represents the statistic for chance performance.

These statistics, as well as the raw data, suggest that localization of overlapping lead and lag bursts of noise in the median plane is more difficult for burst durations less than approximately 10-ms. The statistics further suggest that this increased difficulty is manifested as a decrease in precision and an apparent compression of perceived elevation as well.

B. Discussion

The most notable features in the above-presented single-speaker data are the influence of stimulus duration on source location identification and the compression that characterizes performance for shorter durations. Preliminary experiments (Dizon *et al.*, 1997) indicated that clicks were difficult to localize precisely, while 100-ms noise bursts were well localized. Hence, the present study focused on stimuli with intermediate durations. Evidence of duration-dependent compression in median-plane localization has been reported previously in the literature.

Hofman and Van Opstal (1998) studied median plane localization for stimuli with identical long term spectra, but with varying short term spectra. The intent was to characterize the temporal course of spectral estimation, and also to investigate the ability of subjects to benefit from "multiple looks" of a short-time spectrum. The authors fitted an "elevation gain" to their data, corresponding to the slope of the best fit line to the scatter plot data of response elevation versus actual elevation. Elevation gains indicated a high degree of compression at the shortest durations (3-ms) with increasing gains with longer duration, stabilizing near unity at approximately 80 ms.

Linear regression statistics were calculated on our single-speaker data to permit direct comparison with Hofman and Van Opstal's statistics. Table I lists the correlation coefficient ρ, the elevation gain g, the y-intercept b in degrees, as well as the rms δ of the differences between each data point and the fitted line. The statistics indicate that the correlations are too low for the 1-ms duration for the computed slopes to be meaningful. However, the correlations for 10, 25, and 50-ms are significant. The elevation gains for these durations climb steadily for each subject, indicating less elevation compression as duration is increased from 10-ms $(0.63<g<0.82)$ to 50-ms $(0.91<g<0.95)$. Hofman and Van Opstal's gains, measured for durations between 3 and 80 ms, are quite similar both qualitatively and quantitatively [Table II in Hofman and Van Opstal (1998)].

Macpherson and Middlebrooks (2000) also found elevation compression for short 3-ms bursts that was not evident with the longer 100-ms bursts. However, this result was found to be dependent on presentation level, such that higher level (~50–60 dB SL) stimuli produced significant elevation compression while lower level stimuli did not. The stimuli used in the present study were presented at 40 dB SL only,

TABLE I. Linear regression statistics for the single-speaker data for each subject. Included are the correlation coefficient ρ, the elevation gain g, the y-intercept b in degrees, as well as the rms δ of the differences between each data point and the fitted line.

Subject	Dur (ms)	Corr ρ	Gain g	Offset b (deg)	δ (deg)
S1	1	0.32	0.42	8.58	16
	10	0.82	0.82	0.28	9.8
	25	0.88	0.87	−1.45	8.2
	50	0.92	0.95	−0.39	7.4
S2	1	0.30	0.64	−3.2	25
	10	0.75	0.75	2.4	11
	25	0.86	0.85	1.3	8.6
	50	0.89	0.95	−1.5	8.4
S3	1	0.09	0.24	11	18
	10	0.67	0.64	6.4	11
	25	0.73	0.75	4.1	12
	50	0.89	0.95	1.1	8.7
S4	1	0.09	0.25	−3.1	20
	10	0.49	0.63	1.6	16
	25	0.70	0.79	−2.3	13
	50	0.80	0.91	−1.2	12

which is a level for which the data of Macpherson and Middlebrooks did not show significant compression.

Interestingly, neither Hofman and Van Opstal nor Macpherson and Middlebrooks found any dependence of localization precision on duration, even for a duration of 3-ms, the shortest duration used in either study. In our study, we find a slight increase in variability between 10 and 50-ms (δ in Table I), with a much larger increase in variability for a duration of 1-ms.

The difference in performance for the shortest duration in our study compared with either of the referenced studies may relate to the shorter duration (1-ms as compared to 3-ms), but may also relate to the comb filtering present in our study that was not present in either of the referenced studies. One might expect the comb filtering to influence localization performance due to the additional notches in the high frequency spectrum. However, the comb filtering did not appear to degrade performance for the three longer durations. The elevation gains and the response variability at these durations are quite similar to those found by Hofman and Van Opstal. Perhaps the 1-ms duration combined with the comb filtering was sufficient to degrade performance.

Regarding the compression at the three longer durations, two possible sources for the compression in the elevation responses can be considered: spectral variability in the noise bursts, and edge effects resulting from the response paradigm. Since localization in the median plane is assumed to depend on elevation dependent spectral filtering by the pinnae, it follows that good localization performance is contingent on the subject's ability to acquire an accurate and clean estimate of this spectral filtering (Wightman and Kistler, 1997). Stimuli such as clicks or Gaussian noise are well suited for use in median plane localization experiments due to the smoothness of their magnitude spectra. However, the perceived smooth spectra of Gaussian noise results from peripheral filtering of the finely spaced peaks and notches in the spectra of noise samples of sufficient length. For shorter samples, peaks and notches may persist beyond peripheral filtering due to their wider spacing, and thus may compete with the spectral features that are the primary cues for median plane localization.

In a hypothesized model of performance, Hofman and Van Opstal suggested that the compression observed in their studies could have been due to subjects relying on an initial ‘‘default’’ estimate of elevation when there is an absence of sufficient spectral information to generate a more accurate judgment. In their study, the subject responded with saccadic eye movements, where the saccade always originated from a location directly in front of the listener. In our study, the subject's head was constrained to face directly in front, which is near the center of the speaker array. It is thus possible that in our experiment, subjects were biased toward the center of the distribution when localization was difficult.

Although previous work suggests that elevation compression is not unexpected in these experiments, it is possible that ‘‘edge effects,’’ which refer to the influence of the limited spatial range of response choices available to the listener in this identification task, may contribute to the apparent compression in our study. The two studies discussed above (Hoffman and Van Opstal, 1998; MacPherson and Middlebrooks, 2000) used analog response methods which were effectively unconstrained at the edges.

If we assume here that a listener who localized a stimulus at a position outside the range of responses would choose the speaker nearest the perceived elevation, which would simply be the speaker at the appropriate edge of the response range, we would expect the standard deviation $\sigma[k]$ and the rms error $E[k]$ error to decrease toward the edge speakers since the spatial variance in the responses is rectified by the limited response range. In the data, $\sigma[k]$ is relatively constant across the speaker positions while $E[k]$ grows toward the edge speakers. Thus, the error statistics are not consistent with the hypothesis that the elevation compression is due solely to edge effects. We also note that Hofman and Van Opstal (1998) as well as Macpherson and Middlebrooks (2000) observed compression in elevation using paradigms where edge effects were likely not a factor.

In summary, the most notable features in the data described in the present study are the apparent compression in perceived elevation as well as the increased spatial variance in the responses, especially for the 1-ms duration. It was suggested that the compression may have been partly due to the stimuli themselves (duration dependent spectral variability) or to the response paradigm (edge effects).

IV. PRECEDENCE RESULTS AND DISCUSSION

A. Results

The single-speaker results provide a measure of baseline localization performance of stimuli at the six speaker locations for the four durations chosen. The results with precedence pairs can be compared directly with the single-speaker results as a means of investigating whether and how the location of the lagging noise burst influenced localization. In Fig. 5, results for the PE trials for subject S1 are shown using

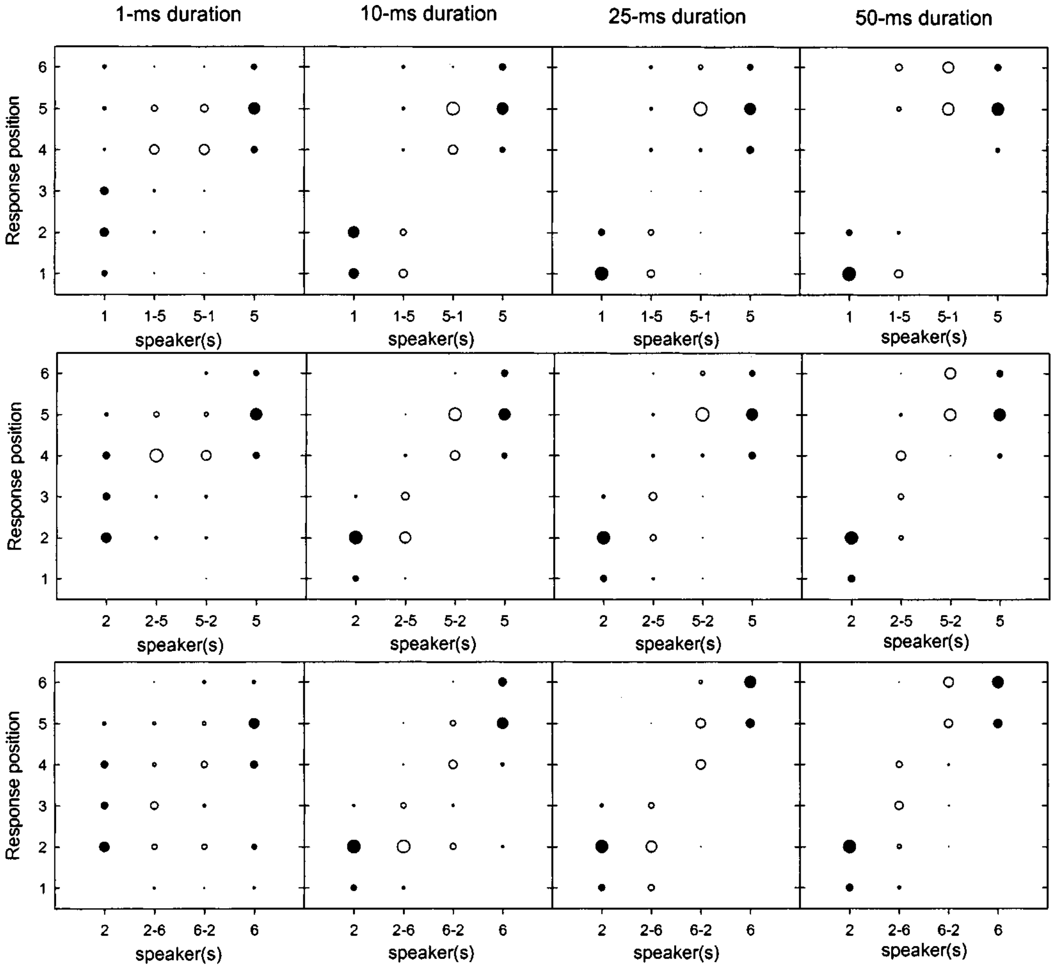

FIG. 5. Raw precedence data for subject S1. Each panel shows response distributions for a position pair and its swapped-speaker counterpart (open symbols), flanked by the single-speaker distributions for each of the speakers in the position pair (closed symbols). Each of the three rows contains data for the same position pair, while the four columns indicate the duration. As in Fig. 3, the area of the bubbles is proportional to the number of responses.

bubble plots similar to those used for the single-speaker data in Fig. 3. Each panel in Fig. 5 shows PE data for a single lead–lag combination, its swapped-pair counterpart, as well as the matching single-speaker results for the lead and lag positions. By comparing the perceived location on single-speaker trials (closed symbols) with that of the PE trials (open symbols) in which either the leading or lagging sources matched the single-speaker position, the extent of PE can be effectively visualized. If the PE is operating, then the distribution of perceived locations on PE trials should be similar to that of the single speaker condition matching the lead. It is important to consider the single-speaker data when viewing the PE data, since a certain amount of the variability and bias in performance should be common to the single-speaker and PE conditions. Certainly, one would not expect performance in the PE trials to be any better than that seen in the single-speaker trials.

In Fig. 5, each row of panels contains data for one lead–lag pair combination, one panel for each of the four durations. The three rows correspond to the three lead–lag pairs chosen (1-5, 2-5, and 2-6), as described in Sec. II. Figures 6–8 show the same data for subjects S2, S3, and S4, respectively.

Considering the data in Figs. 5–8, it should be evident qualitatively that while the localization dominance is not complete, in most conditions there is a clear bias in responses toward the speaker position corresponding to the leading speaker of each PE pair. However, this bias toward the leading speaker is in all cases not as strong as it is for the single-speaker response distribution for the leading speaker. It can be concluded from this difference in bias shifts that localization dominance is not complete, since absolute localization dominance predicts that the precedence stimuli would be localized identically with their single-speaker counterparts corresponding to the leading speaker location. In addition, subjects appear to be able to localize the single-speaker stimuli more precisely, with the exception of the 1-ms duration, where performance is generally poor for both single-speaker and precedence stimuli. Thus, it appears that the presence of the spatially separated lagging stimulus influences the variance in the responses for the combined lead–lag stimulus as well.

The influence of the lag on the shape of the response distributions is often not consistent across subjects for the same position and duration. For instance, for the 50-ms duration, position pair 1-5 response distribution for subject S1 (Fig. 5) is bimodal, while that of the other three subjects are not. For subject S3 (Fig. 7) at a duration of 50-ms, the re-

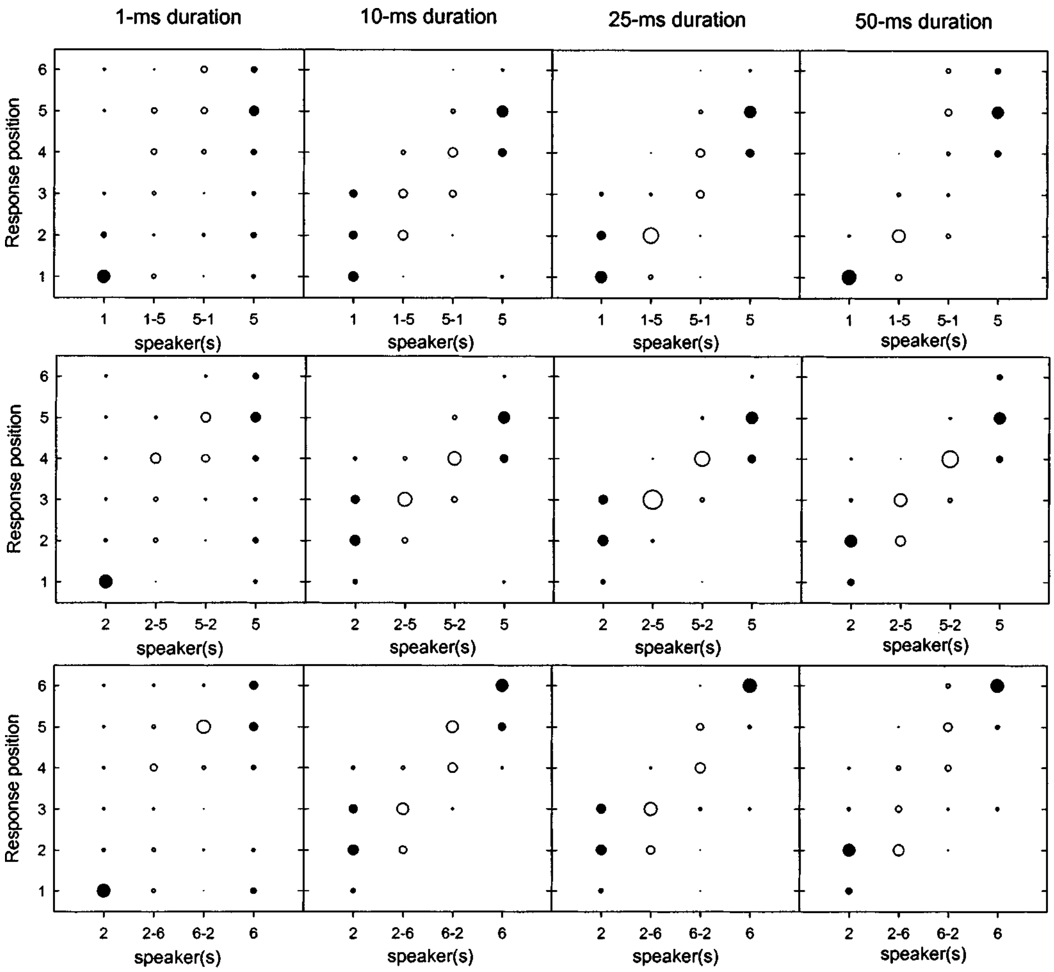

FIG. 6. Same as Fig. 5 but for subject S2.

sponse distribution for position pair 6-2 is clustered very close to position 6, while the response distribution for the swapped-speaker counterpart (2-6) is spread fairly evenly over the whole distribution range. The other three subjects did not exhibit this pattern of performance. Additionally, for the same subject and duration, the response distributions across the three position pairs can differ as well (consider subject S1 at the 50-ms duration).

As a comparison of localization precision between the single-speaker and PE trials, Fig. 9 shows the standard deviations of the response distributions for both single-speaker and PE trials as a function of duration. For each data point, data are averaged over all positions for the single-speaker data, and over all lead–lag pairs for the PE data. Precision appears to be better for the single-speaker trials than for the PE trials at the longer-duration stimuli, as evidenced by the higher standard deviations for the PE trials. Note also that while precision generally improves for single-speaker trials as duration is increased, there is little evidence of an influence of duration on precision for the PE trials, other than an increase in precision in S2's data going from 1 to 10-ms.

The extent of localization dominance in a paired setup can be quantified using a descriptive statistic (accounting for both localization dominance and discrimination suppression) proposed by Shinn-Cunningham (1993) for headphone data and validated for free-field data by Litovsky and Macmillan (1994). A single metric c is calculated, which is bounded between 0 and 1 and represents the extent of leading source localization dominance. It is calculated using the following formula:

$$c=(\alpha_p-\tau_2)/(\tau_1-\tau_2),$$

where α_p is the judged position (or ITD), and τ_1 and τ_2 are the lead and lag positions, respectively. The metric c will equal 1.0 for localization judgments at the leading position, 0.5 for judgments consistent with an equal contribution of lead and lag, and 0.0 for judgements at the lagging position. We extend this model to the median-sagittal plane by using the elevation of the lead and lag speaker positions in degrees for τ_1 and τ_2, and the elevation of the judgment in degrees for α_p.

A c value was computed for every trial at each lead–lag pair. The means are plotted in Fig. 10, with one subject's data in each panel. A thickened horizontal dashed line marking the point at which c equals 0.5 demarcates the boundary between lead dominance and lag dominance in localization. Data for all six of the lead–lag pairs are plotted in a group above the appropriate duration indicated on the x axis. The symbols indicate the lead–lag pair as indicated in the legend, and the error bars represent the standard deviation of c across the 30 trials.

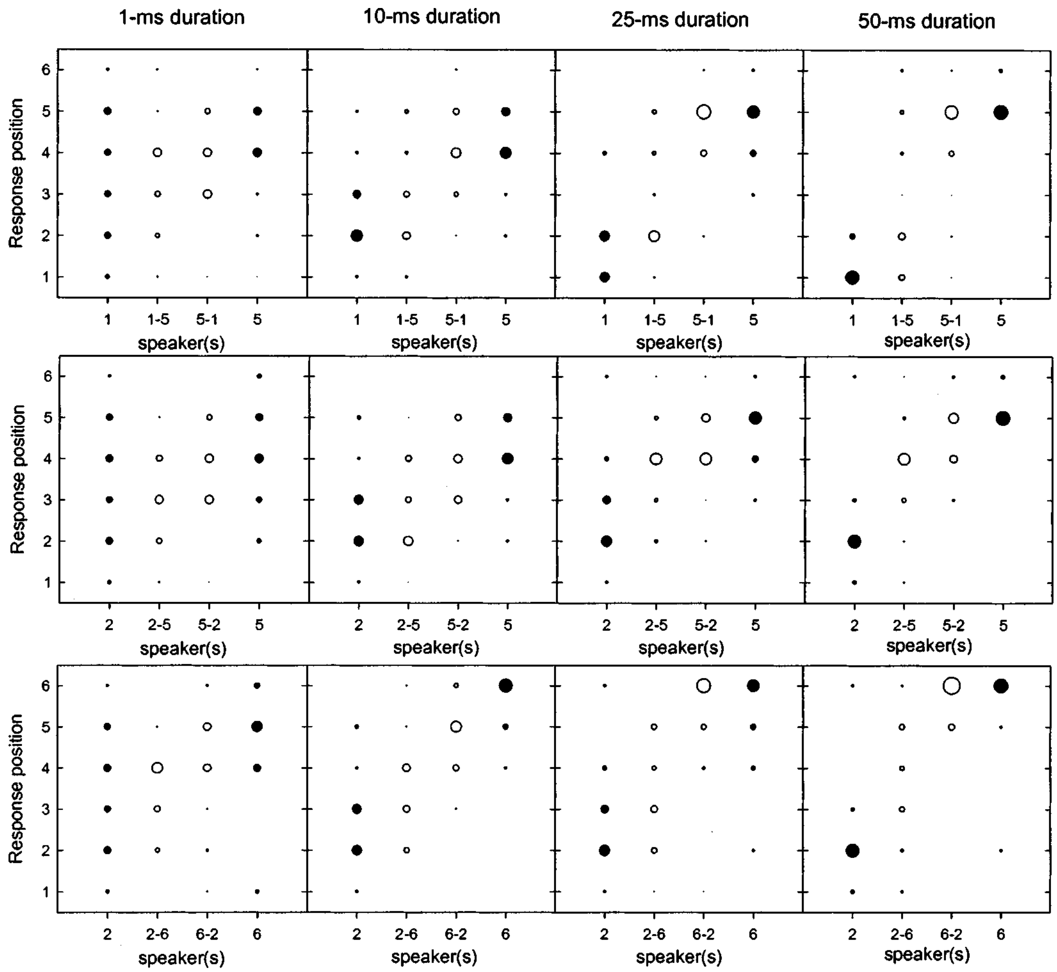

FIG. 7. Same as Fig. 5 but for subject S3.

A majority of the mean values of *c* are above 0.5, with a few close to 1. Note that it is possible for *c* to be above 1 or below 0 if the judgment is not within the range between the lead and lag positions. The standard deviations are quite high, indicating that subjects were not consistent across trials, and frequently localized nearer to the lagging stimulus. These observations on the mean and standard deviations are qualitatively consistent with the scatter plots in Figs. 5–8, which showed response distributions with a lot of spread but generally skewed toward the lead, especially at the longer durations.

Also interesting in Fig. 10 is the variation in *c* over the lead–lag pairs. For example, S3's data at 25 and 50-ms indicate higher *c* values when the lead is at high elevations (the open symbols) as compared to when the lead is low. This may indicate individual bias toward the higher elevations in this paradigm. Conversely, S4's data indicate the opposite, with lower elevations having higher *c* values.

The low means (near 0.5 to 0.7, rather than 1.0) and high variance in the *c* values highlights the fact that the PE is not as strong in the median plane as it is in the free-field azimuthal plane or when ITDs are used over headphones. In the study by Shinn-Cunningham *et al.* (1993) in which *c* values were computed for their own data as well as other studies on precedence using ITDs, *c* values typically reached close to unity (generally above 0.8) for delays near 2 ms. Similarly, Litovsky and Shinn-Cunningham (2001) reported average *c* values of 0.9–1.0 for localization dominance with ITDs at delays of 1 and 2-ms. In free field, Stecker and Hafter (2002) used a similar observer weighting method to estimate localization dominance for lead–lag click pairs, and reported values between 0.7 and 0.8 for delays of 1 to 3-ms.

A note of caution to observing the *c* values alone is that a high *c* value does not necessarily imply localization dominance in our analyses, since a high *c* value (>0.5) for a lead–lag pair combined with a low *c* value (<0.5) for the swapped pair counterpart can indicate a *positional dominance*, signifying a dominance of one elevation over another independent of which contained the leading stimulus. To see why this is the case, recall that when *c* equals zero, the subject localizes toward the lag. If *c* equals one for a lead–lag pair and zero for the swapped pair, then the subject localized both pairs at the same position. Thus, to more effectively gauge the strength of localization dominance, we must look at statistics that compare the response distribution of a lead–lag pair with that of its swapped-pair counterpart.

The following discussion focuses on statistical analyses comparing the responses to presentations of a lead–lag pair with those of its swapped-speaker counterpart. If localization dominance is not effective, then an appropriate statistical test should indicate that the two distributions are not statistically different. The two distributions could have any of the follow-

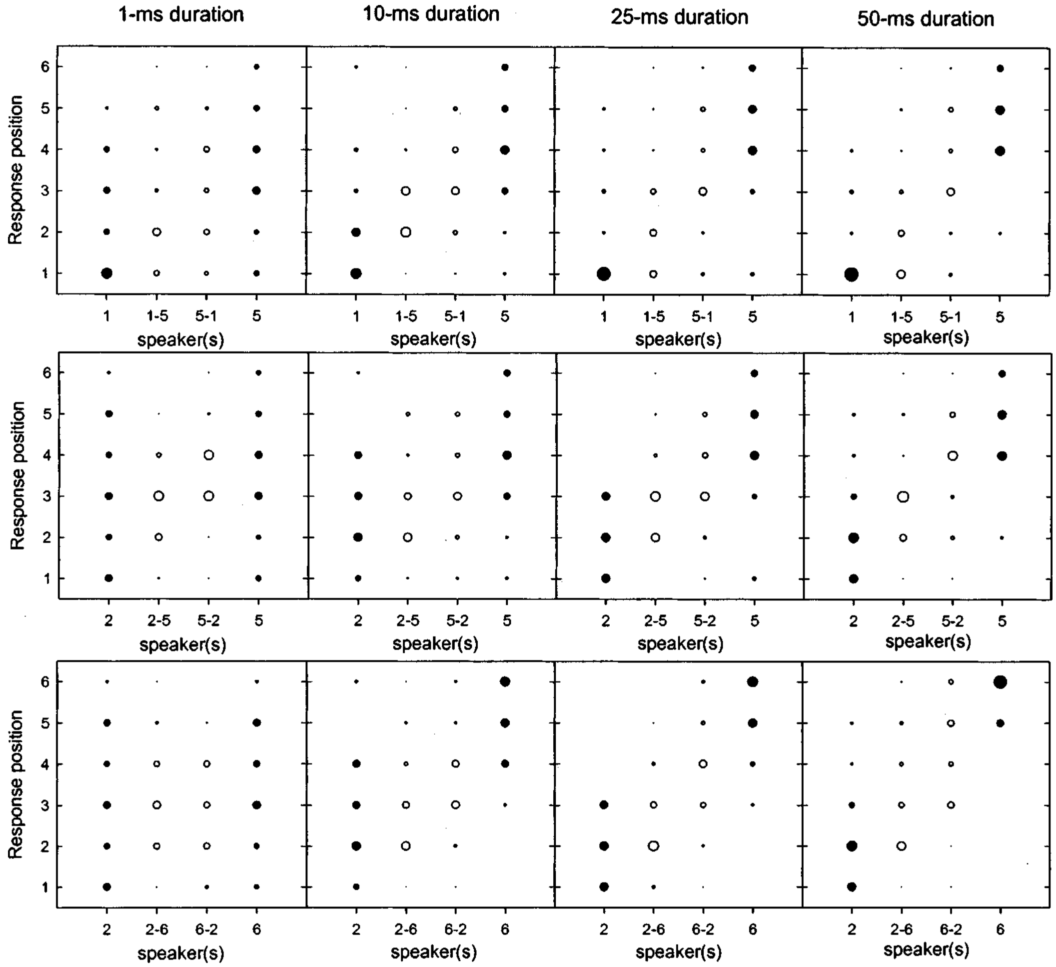

FIG. 8. Same as Fig. 5 but for subject S4.

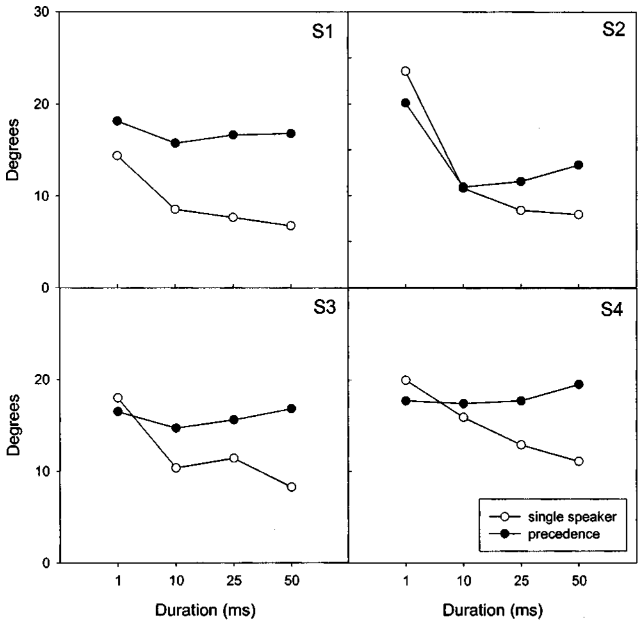

FIG. 9. Standard deviations of the responses for each subject in units of degrees. Each point represents the mean of the individual standard deviations across all positions (for the single-speaker trials) or position pairs (for the precedence trials). Closed symbols represent the precedence trials, open symbols the single-speaker trials.

ing configurations: (1) near one of the two locations, reflecting biased judgments, (2) bimodally distributed in the event that the lead and lag were both heard/localized, or (3) at a location between the lead and lag, in the event that both sounds contributed equally to localization.

A t-test was performed on the response distribution for a given lead–lag pair with that of its swapped-speaker counterpart. For example, for the combinations of 1-5 and 5-1 (stimuli at −30° and +30°), a strong PE would produce distributions that were near location 1 for the 1-5 case and near location 5 for the 5-1 case. The generalized t-statistic for two distributions A and B is given by

$$t_{2N-2}=\frac{M_A-M_B}{\hat{\sigma}_{A,B}},$$

where $\hat{\sigma}_{A,B}$ is a pooled variance given by

$$\hat{\sigma}_{A,B}=\sqrt{\frac{\sigma_A^2+\sigma_B^2}{N}},$$

and where M_A, M_B, σ_A, and σ_B are the mean and standard deviations of the response distributions A and B.

We denote the mean of the distribution for a lead–lag pair as $M[k_1;k_2]$, where k_1 denotes the lead speaker number and k_2 the lag. $M[k_1,k_2]$ is given by

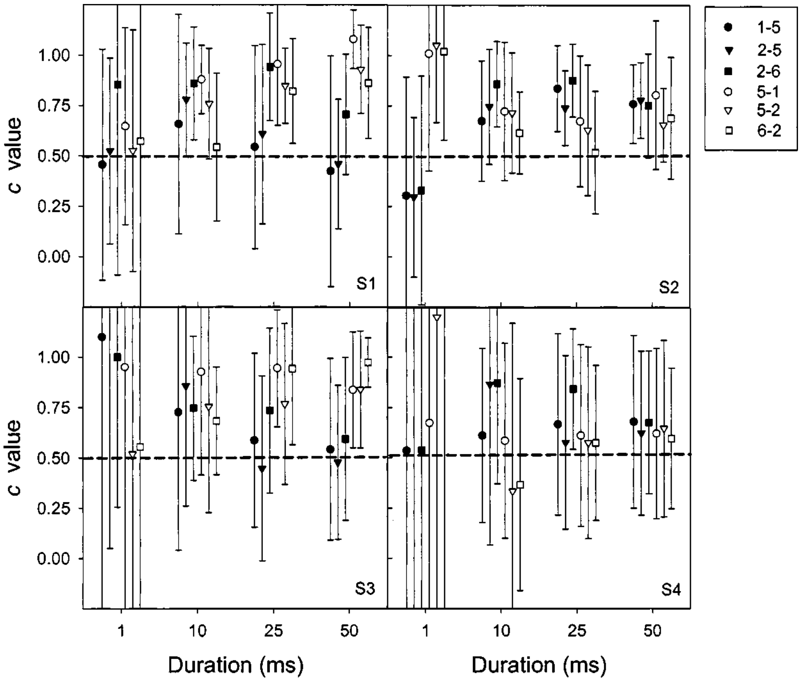

FIG. 10. *c* values computed as indicated in the text. Data represent the mean *c* value across the 30 trials at each position-pair and duration for each subject. Position pair is indicated by symbols defined in the legend. The error bars represent the standard deviation of the *c* across those 30 trials. Data are shown grouped above the duration indicated on the *x* axis, and are shifted laterally for clarity.

$$M[k_1,k_2]=\frac{A}{N}\sum_{n=1}^{N} r_{n,k_1,k_2},$$

where $r_{n,k1,k2}$=response to trial n when the leading sound was presented from speaker k_1 and the lagging sound was presented from speaker k_2. Similarly, $\sigma[k_1;k_2,k_2;k_1]$ denotes the pooled variance for the two swapped-pair distributions, and is given by

$$\hat{\sigma}[k_1;k_2,k_2;k_1]=\sqrt{\frac{\sigma^2[k_1;k_2]+\sigma^2[k_2;k_1]}{N}},$$

where

$$\sigma[k_1,k_2]=\sqrt{\frac{A^2}{N}\sum_{n=1}^{N}(r_{n,k_1,k_2}-M[k_1,k_2])^2}.$$

The t-statistic for the swapped-pair test is given by

$$t_{2N-2}=\frac{M[k_1;k_2]-M[k_2;k_1]}{\hat{\sigma}[k_1;k_2,k_2;k_1]}.$$

Note that this formulation of the t-statistic presumes that $M[k_1;k_2]>M[k_2;k_1]$ for the statistic to be positive. This requires the mean of each distribution to be biased toward the leading side *relative to the other mean*—a requirement that is consistent with localization dominance. This requirement held for all but one of the 48 statistics computed.

It is useful to compare the swapped-pair t-statistic with the "single–single" statistic, which is a t-test comparing the distributions of responses to each position presented in isolation. These distributions are available from the single-speaker trials. If localization dominance was complete, the leading source completely dominates the lag, so we would expect the statistic from the swapped pair comparison to be identical to the single–single statistic.

The t-statistic for the single–single test is given by

$$t_{2N-2}=\frac{M[k_1]-M[k_2]}{\hat{\sigma}[k_2,k_1]},$$

where $\sigma[k_1,k_2]$ is a pooled variance given by

$$\hat{\sigma}[k_2,k_1]=\sqrt{\frac{\sigma^2[k_2]+\sigma^2[k_1]}{N}}.$$

Figure 11 shows the t-statistic for the swapped-pair test on the *y* axis plotted against burst duration on the *x* axis for all four subjects (one in each panel). In each panel, the statistics for each of the three lead–lag pairs (open symbols) are

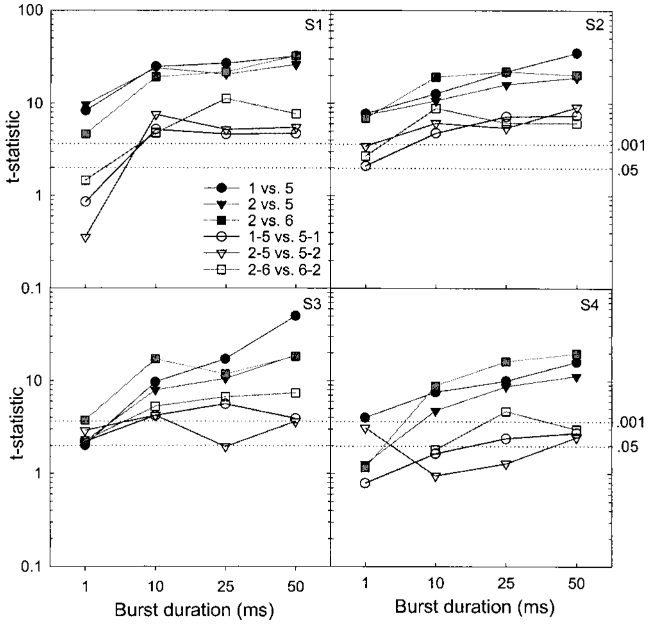

FIG. 11. t-statistics for both the precedence data (open symbols) and the single-speaker data (closed symbols). Symbols indicate position or position pair as shown in the legend. t-statistics for the 0.05 and 0.001 levels are shown as dashed lines.

shown along with those for the single–single comparison (closed symbols). Dashed horizontal lines are shown representing t-statistic values significant to the 0.05 and 0.001 levels. One value is not shown (S4, square symbol at 1-ms) since the means for the two distributions were biased toward the lag rather than the lead. The standard deviations of the response distributions were frequently very high for this stimulus, particularly for subject S4.

The single–single statistics are consistently higher than those for the swapped pairs. This indicates that either the mean shifts are larger for the single–single case, or the variance is smaller. It is notable, however, that the swapped-pair statistics are still highly significant. It is not surprising that the single–single statistics are high given the accuracy and precision of localization of the single-speaker stimuli (Fig. 3). However, the high significance of the swapped pair statistics is not as obvious from a qualitative view of the raw data in Figs. 5–8. The statistical results also suggest some dependence on duration, since the 1-ms statistics are lower for almost all cases. From 10 to 50-ms there is no clear trend in the swapped-pair statistics. It is likely that the higher variability in the responses to these stimuli makes any trends difficult to observe. There is also some inter-subject variation, with subjects S1 and S2 having generally higher statistics overall.

Overall, these analyses indicate that there is a statistically significant temporal order effect favoring the leading stimulus. Recall that this result was not obvious from the *c* values, which consider the distributions in isolation, rather than relative to each distribution's swapped speaker counterpart. Specifically, this means that the response distribution for a lead–lag pair is biased toward the leading stimulus relative to the distribution for its swapped-speaker counterpart. The distributions themselves may be both biased toward one position or another, they may differ in width, or one or both may have bimodal qualities. The raw data in Figs. 5–8 indicate that many of these cases occur. The t-statistics only indicate that the temporal order influences judgments in a manner consistent with localization dominance.

B. Discussion

This study was designed to find evidence for localization dominance in the median-sagittal plane. The two studies cited in Sec. I (Blauert, 1971 and Litovsky *et al.*, 1997a) indicated that certain aspects of precedence exist in the median-sagittal plane, but neither used as fine an elevation spacing as that used in the present study, nor did those studies investigate the influence of duration. The results of the present study are significant in that they are free of front/back confusions (as a result of the positions chosen). Also, they are compared to the single-speaker conditions, allowing the effects of precedence to be separated from localization phenomena related to the comb filtering. This comparison establishes a link between the single-speaker performance and PE performance. Specifically, if a stimulus comprised of a lag co-located with the lead cannot be localized accurately, then one cannot expect to find localization dominance when the lag is at a different location.

The reasons for the poor performance at 1-ms for both the single-speaker and precedence trials could include spectral variability in the short samples of noise, and also the comb filtering imparted by the presence of the lag. As noted in the Single Speaker Discussion (Sec III B), one might have expected the comb filtering to degrade localization for all durations, yet only the 1-ms duration performance is significantly degraded. Perhaps in the 1-ms case, both factors were responsible. The poor single-speaker performance at 1-ms leads logically to the weak evidence of localization dominance at this duration. Specifically, localization dominance of a leading source over a lagging source at a different location is not expected, given that the same stimulus but with the lag co-located with the lead could not be localized precisely.

The choice of positions enabled a better characterization of localization dominance in the median plane than had been shown in the literature. In the present study, both the lead and lag positions are localized accurately, yet the combination produces localization dominance. In the prior median plane studies, there was evidence that a strongly localizable location (front) dominated a weakly localizable location (above) (Litovsky *et al.*, 1997a). This ''positional dominance'' makes it difficult to observe the influence of localization dominance.

The use of longer duration noise bursts in the present study was shown to improve both the accuracy and the precision of single-speaker localization. This performance improvement was helpful for demonstrating statistically significant effects of localization dominance that may not be obvious with shorter duration stimuli. Overall, the results reaffirmed that localization dominance is effective in median plane localization. The results also indicate that the effect is weaker in the median plane as compared to its strength as reported in the azimuthal plane in other studies (e.g., Shinn-Cunningham *et al.*, 1993). Accompanying the present results are considerable inter-subject variability and evidence of positional dependence (variance in performance across position combinations), neither of which are typically as influential in azimuthally based PE experiments (e.g., Litovsky and Shinn-Cunningham, 2001; Saberi and Antonio, 2003).

The reason for weak localization dominance in the median plane is not entirely clear. Echo suppression (measure of whether the lag is heard as a separate sound) appears to be similar in strength in the azimuthal and median-sagittal planes, as measured with single-neuron responses (Litovsky *et al.*, 1997a; Litovsky and Yin, 1998a) or psychophysically (Litovsky *et al.*, 1997b; Rakerd *et al.*, 2000). Hence, it is difficult to argue that auditory mechanisms underlying suppression of echoes are generally weaker in the median plane. However, there is another aspect of the PE, discrimination suppression, which appears to be different in azimuth and elevation. Listeners' ability to discriminate small shifts in the vertical position of the lag is not compromised at any delays, but discrimination of small shifts in the azimuthal position is delay-dependent, being quite poor at brief delays and improving as delays are increased (Litovsky *et al.*, 1997b). Taken together, previous work along with findings from the present study suggest that, while at brief delays the lag may not be subjectively audible, directional properties of the lead do not dominate (take precedence) over those of the lag as strongly in the median plane as they do in azimuth. Weaker dominance of directional cues might be due to the fact that in the median plane there is poorer spatial resolution for source locations than in azimuth. That is, the strength of localization dominance may be a by-product of localization accuracy in each plane.

Computational models of precedence in the azimuthal plane would be difficult to adapt to median plane data. Modeling of elevation localization for single sources is itself not well understood. Computations based on the actual received spectra using the subject's HRTFs as well as the actual noise samples may provide some insight into performance. However, if long term statistics are used, these are likely to bear little fruit in predicting precedence since long-term spectra (assuming that all operations are linear) will be the same if the leading and lagging positions are switched, whereas psychophysical performance clearly depends on which position is leading. More complex models incorporating peripheral nonlinearities (Hartung and Trahiotis, 2001), onset-driven suppression (Lindemann, 1986), or spectral weighting of interaural differences (Tollin, 1998) have shown some success in describing many aspects of azimuthally based precedence. To describe median plane precedence, all of these models would need to be modified to produce an elevation judgment, and may also require sensitivity to spectral profiles. On the other hand, if pinna disparity cues provide a usable cue to median plane localization (e.g., Middlebrooks, 1992; Hofman and Van Opstal, 2003), perhaps interaural difference models of precedence could predict median plane localization dominance as well.

V. CONCLUSIONS

The experiments presented here studied localization dominance in the median-sagittal plane. Localization of single sources in the median plane is mediated primarily by spectral cues, and localization performance in this plane is characterized by front–back confusions as well as high variability when compared to azimuthal localization. The single-speaker data provided a baseline measure of performance in the task using speaker positions in the frontal portion of the median-sagittal plane as a function of burst durations. These data indicated that stimulus duration has a big effect on the single-speaker localization, with longer durations leading to greater precision and less elevation compression. The precedence data were characterized by higher variability in the response distributions, but with a statistically significant bias toward the leading speaker for the longer burst duration conditions. Thus, the *c* values reported here, when compared with those of Shinn-Cunningham *et al.* (1993) are consistent with there being weaker localization dominance in the median-sagittal plane than in the azimuthal plane, not unlike the differences observed by Litovsky *et al.* (1997a).

Considering the precedence effect in general, the existence of localization dominance in the median plane indicates that it is a general phenomenon that applies to paired stimuli in (at least) a two-dimensional space, rather than being confined to the azimuthal dimension only. If one presumes that the phenomenon serves as a means of suppressing reflections, then it should not be surprising that it is a general spatial effect.

ACKNOWLEDGMENTS

The authors are extremely grateful to Dr. H. S. Colburn for numerous helpful discussions of the data and comments on previous versions of the manuscript, and to Gerald Ng for help and participation in earlier versions of these studies. This work was supported by NIH-NIDCD Grant Nos. R01-DC003083 and R01-00100. Portions of the data were presented at the 133rd and 139th Meetings of the Acoustical Society of America.

Blauert, J. (**1971**). ‘‘Localization and the law of the first wavefront in the median plane,’’ J. Acoust. Soc. Am. **50**, 466–470.

Blauert, J. (**1997**). *Spatial Hearing: The Psychophysics of Human Sound Localization*, Revised ed. (MIT, Cambridge, MA).

Butler, R. A. (**1969**). ‘‘Monaural and binaural localization of noise bursts vertically in the median sagittal plane,’’ J. Aud Res. **3**, 230–235.

Dizon, R., Litovsky, R. Y., and Colburn, H. S. (**1997**). ‘‘Positional dependence on localization dominance in the median-sagittal plane,’’ J. Acoust. Soc. Am. **101**, S3106.

Fitzpatrick, D. C., Kuwada, S., Batra, R., and Trahiotis, C. (**1995**). ‘‘Neural responses to simple, simulated echoes in the auditory brainstem of the unanesthetized rabbit,’’ J. Neurophysiol. **74**, 2469–2486.

Gardner, M. B., and Gardner, R. S. (**1973**). ‘‘Problem of localization in the median plane: Effect of pinnae cavity occlusion,’’ J. Acoust. Soc. Am. **53**, 400–408.

Hartmann, W. M. (**1983**). ‘‘Localization of sound in rooms,’’ J. Acoust. Soc. Am. **74**, 1380–1391.

Hartung, K., and Trahiotis, C. (**2001**). ‘‘Peripheral auditory processing and investigations of the ‘precedence effect’ which utilize successive transient stimuli,’’ J. Acoust. Soc. Am. **110**, 1505–1513.

Hebrank, J., and Wright, D. (**1974**). ‘‘Spectral cues used in the localization of sound sources on the median plane,’’ J. Acoust. Soc. Am. **56**, 1829–1834.

Hofman, P. M., and Van Opstal, A. J. (**1998**). ‘‘Spectro-temporal factors in two-dimensional human sound localization,’’ J. Acoust. Soc. Am. **103**, 2634–2648.

Hofman, P. M., and Van Opstal, A. J. (**2003**). ‘‘Binaural weighting of pinna cues in human sound localization,’’ Exp. Brain Res. **148**, 458–470.

Lindemann, W. (**1986**). ‘‘Extension of a binaural cross-correlation model by contralateral inhibition. II. The law of the first wavefront,’’ J. Acoust. Soc. Am. **80**, 1623–1630.

Litovsky, R. Y., Colburn, H. S., Yost, W. A., and Guzman, S. (**1999**). ‘‘The precedence effect,’’ J. Acoust. Soc. Am. **106**, 1633–1654.

Litovsky, R. Y., Hawley, M. L., Dizon, R., and Colburn, H. S. (**1997b**). ‘‘Measurements of precedence phenomena in binaural and monaural conditions,’’ J. Acoust. Soc. Am. **101**, S3083.

Litovsky, R., and Macmillan, N. (**1994**). ‘‘Minimum auditory angle for clicks with simulated echoes: Effects of azimuth and standard,’’ J. Acoust. Soc. Am. **96**, 752–758.

Litovsky, R. Y., and Yin, T. C. T. (**1998a**). ‘‘Physiological studies of the precedence effect in the inferior colliculus of the cat. I. Correlates of psychophysics,’’ J. Neurophysiol. **80**, 1285–1301.

Litovsky, R. Y., and Yin, T. C. T. (**1998b**). ‘‘Physiological studies of the precedence effect in the inferior colliculus of the cat. II. Neural Mechanisms,’’ J. Neurophysiol. **80**, 1302–1316.

Litovsky, R. Y., Yin, T. C. T., Rakerd, B., and Hartmann, W. M. (**1997a**). ‘‘Psychophysical and physiological evidence for a precedence effect in the median sagittal plane,’’ J. Neurophys. Rapid Communication **77**, 2223–2226.

Litovsky, R. Y. and Shinn-Cunningham, B. G. (**2001**). ‘‘Investigation of the relationship among three common measures of precedence: fusion, localization dominance, and discrimination suppression,’’ J. Acoust. Soc. Am. **109**, 346–358.

Macpherson, E. A., and Middlebrooks, J. C. (**2000**). ‘‘Localization of brief sounds: Effects of level and background noise,'' J. Acoust. Soc. Am. **108**, 1834–1849.

Middlebrooks, J. C. (**1992**). ''Narrow-band sound localization related to external ear acoustics,'' J. Acoust. Soc. Am. **92**, 2607–2607.

Middlebrooks, J. C., and Green, D. M. (**1991**). ''Sound localization by human listeners,'' Annu. Rev. Psychol. **42**, 135–159.

Rakerd, B., Hartman, W. M., and Hsu, J. (**2000**). ''Echo suppression in the horizontal and median sagittal planes,'' J. Acoust. Soc. Am. **107**, 1061–1064.

Saberi, K. and Antonio, J. V. (**2003**). ''Precedence-effect thresholds for a population of untrained listeners as a function of stimulus intensity and interclick interval,'' J. Acoust. Soc. Am. **114**, 420–429.

Searle, C. L., Braida, L. D., Cuddy, D. R., and Davis, M. F. (**1975**). ''Binaural pinna disparity: Another auditory localization cue,'' J. Acoust. Soc. Am. **57**, 448–455.

Shinn-Cunningham, B. G., Zurek, P. M., and Durlach, N. I. (**1993**). ''Adjustment and discrimination measurements of the precedence effect,'' J. Acoust. Soc. Am. **93**, 2923–2932.

Stecker, G. C. and Hafter, E. R. (**2002**). ''Temporal weighting in sound localization,'' J. Acoust. Soc. Am. **112**, 1046–1057.

Tollin, D. J. (**1998**). ''Computational model of the lateralisation of clicks and their echoes,'' in ''Proceedings of the NATO Advanced Study Institute on Computational Hearing,'' edited by S. Greenberg and M. Slaney, I-Lucca, Italy, 1–12 July, 1998, pp. 77–82.

Wallach, H., Newman, E. B., and Rosenzweig, M. R. (**1949**). ''The precedence effect in sound localization,'' Am. J. Psychol. **LXII**, 315–336.

Wightman, F. L., and Kistler, D. J. (1997). ''Factors affecting the relative salience of sound localization cues,'' in *Binaural and Spatial Hearing in Real and Virtual Environments*, edited by R. H. Gilkey and T. R. Anderson (Earlbaum, Mahwah, NJ).

Yin, T. C. T. (**1994**). ''Physiological correlates of the precedence effect and summing localization in the inferior colliculus of the cat,'' J. Neurosci. **14**, 5170–5186.

Zurek, P. M. (**1980**). ''The precedence effect and its possible role in the avoidance of interaural ambiguities,'' J. Acoust. Soc. Am. **67**, 952–964.