The precedence effect

Ruth Y. Litovsky[a] and H. Steven Colburn
Hearing Research Center and Department of Biomedical Engineering, Boston University, Boston, Massachusetts 02215

William A. Yost and Sandra J. Guzman
Parmly Hearing Institute, Loyola University Chicago, Chicago, Illinois 60201

(Received 20 April 1998; revised 9 April 1999; accepted 23 June 1999)

In a reverberant environment, sounds reach the ears through several paths. Although the direct sound is followed by multiple reflections, which would be audible in isolation, the first-arriving wavefront dominates many aspects of perception. The ''precedence effect'' refers to a group of phenomena that are thought to be involved in resolving competition for perception and localization between a direct sound and a reflection. This article is divided into five major sections. First, it begins with a review of recent work on psychoacoustics, which divides the phenomena into measurements of fusion, localization dominance, and discrimination suppression. Second, buildup of precedence and breakdown of precedence are discussed. Third measurements in several animal species, developmental changes in humans, and animal studies are described. Fourth, recent physiological measurements that might be helpful in providing a fuller understanding of precedence effects are reviewed. Fifth, a number of psychophysical models are described which illustrate fundamentally different approaches and have distinct advantages and disadvantages. The purpose of this review is to provide a framework within which to describe the effects of precedence and to help in the integration of data from both psychophysical and physiological experiments. It is probably only through the combined efforts of these fields that a full theory of precedence will evolve and useful models will be developed.
[S0001-4966(99)01910-4]

PACS numbers: 43.10.Ln, 43.71.An, 43.71.Bp [ADP]

INTRODUCTION

This paper reviews recent work and current thinking about a group of auditory phenomena that are thought to account for listeners' abilities to function in reverberant spaces, and that for historical reasons have been attributed to the ''precedence effect.'' The most extensive review to date on this topic is in Blauert's classic book on spatial hearing which was just recently expanded and reprinted (Blauert, 1997). Blauert reviews the classic data on precedence up to about 1982, and then in a newly added chapter he reviews some of the recent work on the buildup phenomenon (which is covered in Sec. II of this paper). Zurek (1987) also provides a review of the work through the mid-1980s, and related chapters can be found in a recent book edited by Gilkey and Anderson (1997). In recent years there has been a resurgence of interest in the precedence effect by psychoacousticians and physiologists, and a new body of literature has been produced which has led us, and many others, to reevaluate our assumptions about the auditory and neural mechanisms involved. This review is our attempt to encompass all of these topics, especially those that have been studied since the mid-1980s.

When a sound is produced in a reverberent environment, it propagates in multiple directions and is subsequently reflected from nearby surfaces. The auditory system is thus faced with resolving competition between the first sound and its reflections for perception and localization. Despite this clutter of information, we can localize sound sources and identify their meaning fairly accurately. Figure 1 illustrates a recording of a source click and its reflections. The recording was made in a ''typical'' classroom (approximately 6 $\times 11$ m). The sound source (S) was a brief click delivered 1.3 m in front of a blackboard and the recording measurement was made 4 ft in front of the source. Three reflections can be identified and these are labeled $R1$, $R2$, and $R3$. They occur approximately 8, 8.5, and 10 ms after the direct click and are attenuated relative to the source. For example, the first reflection occurs about 8 ms after the source and its level is attenuated by about 9.5 dB from the level of the source click. These reflections are sometimes referred to as ''early reflections'' to differentiate them from the total reverberation created by the interaction of all the reflections (i.e., the total acoustic clutter produced by a sound in a reflective environment). In general, a reflection is an attenuated, sometimes spatially separated, delayed and coherent copy of the originating sound.

As a simplification of a natural situation, consider an arrangement of two loudspeakers in an anechoic room such that the speakers are equally distant, and stimulated by identical sounds such that the onset of one sound is delayed relative to the onset of the other sound. This can be considered a model of a direct sound (the lead) with a single reflection (the lag). This situation for click stimuli is shown in Fig. 2(A), and an idealized sketch of the perceived location(s) of the image(s) as a function of the lead–lag delay is shown in Fig. 2(B). The lead is at 45° to the right, and the lag is at 45° to the left. When the delay is zero and the speakers are stimulated equally, the stimuli to the two ears of the

[a]Address for correspondence: Boston University, Biomedical Engineering, 44 Cummington St., Boston, MA 02215; electronic mail: litovsky@bu.edu

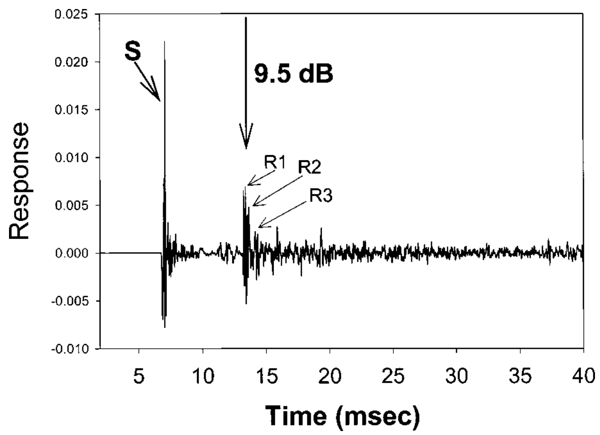

FIG. 1. The recording of the response to a 10-μs electrical pulse in a college classroom that was approximately 6×11 m in size. The source of the transient was placed 1.3 m in front of a blackboard attached to the front of the classroom and the recording was made 4 ft in front of the source. The first 40 ms of the recording shows the direct click (S) and several early reflections labeled as $R1$, $R2$, and $R3$. These reflections stand out from the overall reverberation.

listener are approximately equal and a single (fused) image is perceived in the plane of symmetry, approximately straight ahead of the listener. As the delay increases, the fused image moves toward the direction of the lead speaker, reaching this direction after about a millisecond. For delays between about 1 and 5 ms, the image is still unitary (fused) and remains located in the direction of the leading speaker. Finally, for large delays, the image breaks into two images, one at each

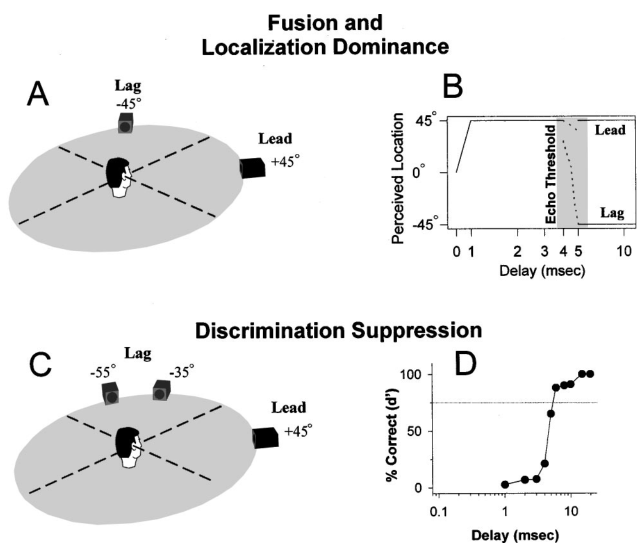

FIG. 2. Schematic diagram of spatio-temporal relation for a lead–lag click-pair. Panels A and C give examples of stimulus configurations for the fusion/localization and discrimination tasks, respectively. In panel A the lead is at 45° to the right, and the lag is at 45° to the left. Panel B shows changes in perceived locations of the auditory events as a function of the delay. With no delay, one fused image is heard at a ''phantom'' location between the lead and lag. Between 0–1 ms the image shifts toward the lead speaker. Between 1 ms and the ''echo threshold'' a fused image is heard at the lead location. When echo threshold is reached, a second image appears, initially near the lead location, and at longer delays at the lag location. In C the lead is at 45° to the right, and the two lag locations are at 35 and 55° to the left. Panel D shows sample data for the discrimination task, in which performance is poor at short delays and improves as delays are increased (panel B is modified from Blauert, 1997, with permission).

speaker position, as one would expect for sequential, well-separated stimuli.

This idealized situation can be used to provide the vocabulary used to describe these phenomena in general. ''Summing localization'' refers to a delay (0–1 ms) when the sounds from the lead and lag sources are perceptually fused and when both the lead and lag contribute to the perceived direction of the fused image (e.g., de Boer, 1940; Warncke, 1941; for review see Blauert, 1997, pp. 204–206). Note that the simplest case of summing localization, as illustrated in Fig. 2(B), assumes no temporal overlap between the direct and reflected signals and the perceived location is an average of the two directions. In cases where the stimuli overlap in time, perceived direction is mediated by more complex averaging that include the amplitudes and phases of the summed wave forms.

As the delays are increased beyond 1 ms several observations can be made. At relatively short delays (typically in the range of 1 to 5 ms, or more, depending on the stimulus wave form and room acoustics) the two sounds remain perceptually fused, hence we refer to this percept as ''fusion.'' As the delay increases, the lagging source becomes audible as a separate auditory event; this perceptual boundary between ''one fused sound'' and ''two separate sounds'' is often referred to as the echo threshold. Blauert (1997) emphasizes this definition for two spatially separated sources.

We emphasize the fact that echo threshold is not the threshold of detectability of the lag; lead–lag sounds and lead-only sounds can be distinguished easily based on overall sound quality (e.g., Blauert, 1997). The echo threshold estimates the delay at which the fused image perceptually splits into two images. This distinction is important since information contained in reflections is important for our ability to sense that the environment we are in is reverberant.

In addition to fusion, the finding that at short delays the image location is dominated by the location of the leading source has been called the ''law of the first wave front'' or ''localization dominance.'' Finally, ''lag discrimination suppression'' refers to the observation that at short delays stimulus parameters of the lag stimulus are less discriminable due to the presence of the lead stimulus. Discrimination improves as the delays increase [see Fig. 2(C) and (D)]. In part, the purpose of this review is to introduce the vocabulary of fusion, localization dominance, and lag discrimination suppression as a means to organize the various ''effects'' associated with what has been called the precedence effect.

Although the lead–lag stimulus paradigm is quite idealized compared with realistic stimuli in reverberant spaces, it has become widely used in psychophysical and (in recent years also in) physiological studies. The term ''precedence effect'' was originally coined by Wallach *et al.* (1949) in their classic study to describe the dominance of the lead stimulus characteristics in the determination of the spatial location of the fused image (localization dominance). However, in the past two decades this term has become popularized and is used to refer to most measurements made using lead–lag stimulus configurations, regardless of the psychophysical measurement that is made. Hence, fusion, localization dominance, and discrimination suppression have all been used somewhat interchangeably. We view this as a potential problem, which can easily result in confusion when trying to understand the mechanisms underlying these perceptual phenomena. The primary goal of this paper is to carefully delineate between studies that claim to measure ''the precedence effect,'' but which may measure different perceptual effects. When appropriate, we will attempt to draw parallels and to link related findings in these areas. To date, few studies exist in which all three measurements have been made in the same listeners, hence to the extent to which fusion, localization dominance, and discrimination suppression are directly related is not clearly understood.

In addition to delineating between perceptual phenomena, we feel that it is important to keep in mind which method of stimulus presentation is used in the various studies. The situation depicted in Fig. 2 represents aspects of the precedence effect in a free-field environment. However, much of what is known comes from headphone studies, in which stimuli from loudspeakers at different locations are replaced by stimuli with different interaural time or level differences for the lead and lag sources (e.g., Wallach *et al.*, 1949; Zurek, 1980; Gaskell, 1983; Yost and Soderquist, 1984; Shinn-Cunningham *et al.*, 1993). There have also been a few experiments with headphones using virtual acoustic stimuli that attempt to recreate the same stimuli in the ear canals that arise from free-field stimulation (e.g., Dizon *et al.*, 1997; Litovsky *et al.*, 1999). These experiments allow manipulation of spectral, temporal, and level differences separately with the simulated free-field condition as the reference condition. We draw attention to this distinction because some measurements indicate that effects differ depending on the method of stimulus presentation. In Fig. 3, the stimulus configuration for the free field is shown in panel A, where the lead and lag are each shown as arriving from different locations with a delay between their onsets. The acoustic signals at the left and right earphones for the headphone stimulation case are shown in panel B, where the lead and lag each contain interaural time differences that result in images with different perceived lateralization. Finally, the resulting wave forms at the ears for the free-field case are sketched in panel C.

It is clear from Fig. 3 that stimuli that are often used in studies of the precedence effect differ dramatically from ''realistic'' stimuli. The lead/lag simulation of a sound source and its reflection differs in several ways from the acoustics in real rooms. Lead and lag stimuli are often clicks of equal amplitude and identical wave forms, while sound sources need not be transient and reflections are usually different from the sound source in amplitude and wave form. Nonetheless, click stimuli have been popular in precedence studies due to their transient nature, which avoids temporal overlap between the lead and lag. About 50 years ago there were some very interesting studies on the processing of complex stimuli in a source-reflection paradigm (e.g., Haas, 1949), and that work is reviewed in detail by Blauert (1997). We believe that studies on precedence, in which the stimuli are optimized and nonrealistic, represent the first step toward understanding basic auditory processes that are involved in resolving competition for perception and localization between a source and its reflection. Ultimately, experimenters should aim to use more realistic stimuli; however, while the mechanisms underlying precedence are not well understood, minimizing stimulus complexity remains essential.

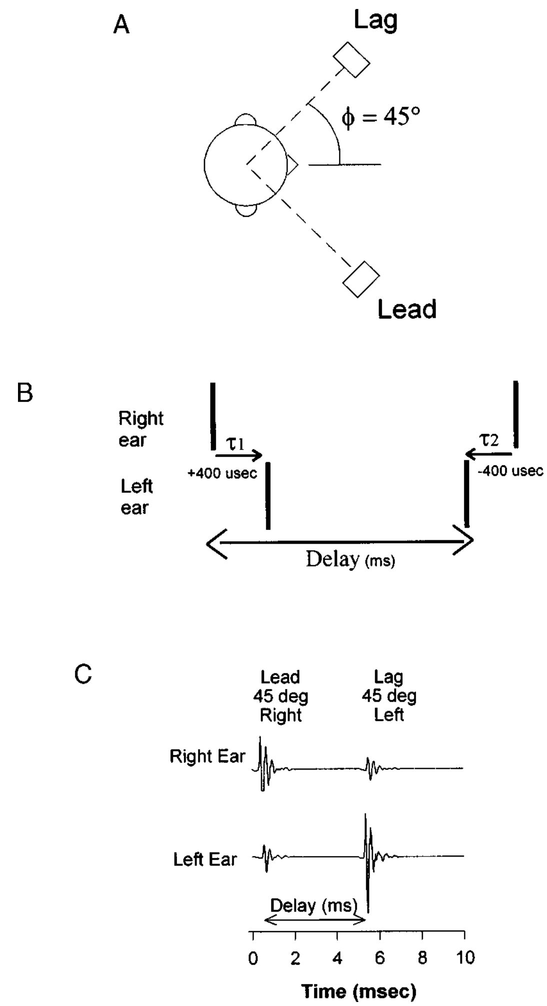

FIG. 3. Stimulus configurations commonly used in precedence studies. A: Free-field setup. Two sounds are emitted from locations 45° to the right and left with the right signal leading the left signal by several milliseconds. B: Dichotic headphone stimuli used by Wallach *et al.* (1949). Two click pairs with ITDs leading to the right ear and then the left ear, with a delay of a few milliseconds between the pairs. The individual wave forms at each ear for the lead and the lag pairs would be approximately the same levels and differ only in the interaural delay. C: Sketch of two impulse responses to click-pairs as the sound sources reach the ears in free field from 45° to the right and left. There are natural interaural differences in time and intensity between the stimuli at the two ears, as well as some differences in the spectral shape of the sounds. Finally, to simulate the precedence effect there is an interstimulus delay, simulating the echo delay.

In this paper we will use lead and lag stimuli to refer to conditions involving the types of simulations we have just described. We will reserve the terms ''sound source'' and ''reflections'' to discussions of real acoustic environments or applications of experimental work to the real world. We also will reserve the word ''echo'' for situations in which fusion breaks down and a reflection or lag stimulus is perceived as a separate sound source. The vocabulary we are applying to the study of precedence in this review is not necessary for understanding the literature. However, we believe that it will help the reader *better* understand the literature, and it will

TABLE I. Crticial thresholds for fusion, discrimination suppression, and localization dominance.

Study	Stimulus	Thresholds	Criterion for threshold
FUSION ECHO THRESHOLDS			
Haas (1951)	speech	30–40 ms	''echo annoying''
Lochner and Burger (1958)	speech	50 ms	lead and lag ''equally loud''
Schubert & Wernick (1969)	noise		
	a) 20-ms duration	5–6 ms	lead and lag ''equally loud''
	b) 50-ms duration	12 ms	
	c)100-ms duration	22 ms	
Ebata *et al.* (1968)	clicks	10 ms	fused image at center of the head
Freyman *et al.* (1991)	clicks	5–9 ms	lag heard on 50% of trials
Yang and Grantham (1997a)	clicks	5–10 ms	lag clearly audible on 75% of trials
Litovsky *et al.* (1999)	clicks	5–10 ms	lag clearly audible on 75% of trials
DISCRIMINATION CRITICAL THRESHOLDS			
Freyman *et al.* (1991)	clicks	5–9 ms	$d' = 1$
Yang and Grantham (1997b)	clicks	5–10 ms	discrimination 75% correct
Litovsky *et al.* (1999)	clicks	5–10 ms	discrimination 75% correct
LOCALIZATION CRITICAL THRESHOLDS			
Litovsky *et al.* (1997b)	clicks	11.4 ms	lead location chosen on 75% of trials
Litovsky *et al.* (1997a)	clicks	8 ms	lead location chosen on 75% of trials

facilitate the integration of information about precedence as well as aid in modeling sound processing in reflective environments.

Most of this review is concerned with recent psychophysical research related to the precedence effect; this constitutes the material of Sec. I. In separate subsections, measurements of fusion, localization dominance, and discrimination suppression are discussed individually. Section II considers recent results showing how preceding stimulation affects perception of simulated reflections, and we describe two important findings based on repeating lead–lag stimulus pairs: the buildup and the breakdown of some of the effects of precedence. In Sec. III, material related to human development and animal experiments is considered. In Sec. IV, attention is given to neural mechanisms that might underlie psychophysical phenomena related to precedence; hence, we will attempt to weave together and correlate findings in psychophysics and physiology. Finally, in Sec. V we consider current models of precedence and their limitations.

I. PSYCHOPHYSICS WITH SIMPLE PAIRED STIMULI

A. Fusion

Fusion is a striking perceptual effect: at short delays between a sound source and its reflection, while two or more equally loud, spatially separated sounds are physically present, listeners perceive only one fused auditory event. This effect can be quite useful for avoiding multiple sound images that may arise from the source and its reflections. A listener's ability to negotiate sounds in a reflective space can be enhanced if the auditory system groups a sound source and its reflections in to a single, coherent auditory percept.

In order to measure fusion in a controlled laboratory environment, listeners are usually asked to provide a subjective impression of how many sounds they hear, and measurements are usually repeated at various lead–lag delays. Results can be plotted as psychometric functions that show the percent of trials on which ''two sounds'' are reported as a function of delay. At short delays (<5 ms for clicks) most listeners report hearing one sound on the majority of trials, and as the delays are increased the proportion of trials on which ''two sounds'' are reported increases, usually reaching 100% by 8 to 10 ms. In addition to the audibility of the lag, other perceptual changes occur, including the influence of the lag on aspects of the fused image, such as its loudness, spatial extent, and pitch. These types of perceptual changes also depend on the type of signal, the signal level, the direction of the sources, and whether the sounds are presented in free field or over headphones. It is important to note that the detectability of the lagging source is not of primary interest; listeners are usually able to distinguish between trials in which a lag is present and trials in which the lead is presented alone (Guzman and Yost, 1999). Rather, the focus is on whether the lag is perceptually fused with the lead or whether it is perceived as a separate sound.

A common measure of the temporal boundary between separating perception of ''one fused sound'' from ''two separate sounds'' is the *echo threshold* (for review see Zurek, 1987; Blauert, 1997). Quantitative estimates of echo thresholds vary tremendously (2–50 ms), depending on a number of variables. The primary determinant of the echo threshold seems to be the nature of the stimulus, although instructions to subjects are probably important and few laboratories used several types of stimuli within a single study. In addition, spatial separation might be a determinant. Blauert's (1997) definition assumes that the lead and lag are separated while not all other writings do. In Table I we list examples of echo thresholds obtained with different stimuli, showing that echo thresholds are much shorter for brief stimuli such as clicks than for longer duration stimuli such as noise and ongoing complex stimuli such as running speech. Although the table is not inclusive of all studies conducted on fusion, it represents a good summary of the stimulus variables that might influence echo threshold.

Fusion studies have an inherent subjective nature, hence estimates of echo threshold also depend on the instructions given to the listener. Thus echo thresholds can vary depending on whether one measures delays at which the ''primary auditory event and reflection [are] equally loud'' (e.g., Haas, 1951; Meyer and Schrodder, 1952; Lochner and Burger, 1958; Franssen, 1960, 1963; David, 1959), the ''reflection [is] annoying'' (Haas, 1951; Muncey *et al.*, 1953), ''a second sound is heard at the vicinity of the lag speaker'' (e.g., Freyman *et al.*, 1991), or ''one or more than one potential sources of the perceived sound are detected'' (e.g., Yost and Guzman, 1996). A large majority of studies on fusion was conducted prior to 1971 and usually in the free field. These studies are discussed extensively by Blauert (1997). Much less work has been done over headphones, and this work (e.g., Yost and Soderquist, 1984) suggests that echo thresholds are shorter for headphones than for free-field delivered stimuli.

More recently, several additional aspects of fusion have been noted. There seems to be significant intersubject variability in the strength of fusion (e.g., Clifton and Freyman, 1989; Freyman *et al.*, 1991; Clifton *et al.*, 1994). With click stimuli, some listeners no longer experience fusion at relatively short delays (2–4 ms), while others experience a strong effect that lasts beyond 10 ms. In addition, it has been suggested (Litovsky and Colburn, 1998) that spatial separation between lead and lag significantly reduces echo threshold.

In the classic literature on precedence there seems to be an underlying assumption that the precedence effect is a binaural phenomenon, mediated by the binaural circuits in the auditory system, and hence it has been modeled using binaural inputs (e.g., Blauert, 1997). It turns out that several aspects of precedence, including fusion, occur at similar delays under binaural and monaural conditions.

One might ask what natural situation would result in presentation of sounds via a monaural system? In recent years at least two such scenarios have been identified and studied. The first has been to compare listeners' performance in the azimuthal and median-sagittal planes, where the relative strength of binaural disparity cues and spectral cues differ in the two dimensions (Litovsky *et al.*, 1999; Rakerd *et al.*, 1997). The second approach has been to measure fusion in listeners who have profound monaural deafness (Litovsky *et al.*, 1997a). Monaural listeners are accustomed to functioning in their environment using information from one ear alone, hence they probably represent a ''perfect'' monaural system much more realistically than would normal-hearing listeners tested under monaural conditions. The results are suggestive of a fusion mechanism that is not dependent on binaural processing. In the azimuthal-median studies, the delay at which listeners perceived ''one sound'' or ''two sounds'' for either clicks (Litovsky *et al.*, 1999) or speech (Rakerd *et al.*, 1997) is nearly identical for most listeners. In the study using monaural listeners, fusion was found at similar delays for the normal-hearing and the monaural populations.

To summarize, a basic perceptual consequence of listening in reflective environments is that sources and their reflections become perceptually fused into a single coherent auditory percept. This effect is strongest at short delays (1–8 ms), not unlike those that occur in relatively small enclosed spaces. Sound travels at the speed of approximately 1 meter per 3 ms, hence a reflection might take up to 6 ms to reach the ears from a nearby wall, and up to 10 ms or longer from a far wall. The advantages of fusion might therefore be most noticeable in small rooms, and decrease as the reflective surfaces are placed farther from the listener. Fusion studies have been mostly limited to subjective impressions reported by listeners, with little emphasis on the perceived location of the lead, lag, or fused image. The next section on localization dominance focuses on processing of directional information.

B. Localization dominance

It is not hard to imagine what a listener might experience if directional information contained in a source and in its reflections were weighted equally by the auditory system. It would be difficult to identify the true location of the source. However, that is not the case. The reader can relate to his/her own experience, whereby a source can usually be correctly localized, and the reflections contribute relatively little directional information. This everyday experience is what we term ''localization dominance.'' In a controlled laboratory situation this phenomenon is usually studied by simulating one source (lead) and one reflection (lag), and the effect is thought to be strongest at short delays (greater than a millisecond and shorter than echo threshold). This is not to say that directional information from the lag is completely ignored, but that the contribution of the lead to localization of the fused image is much stronger than that of the lag.

1. Headphone studies

The bulk of studies on localization dominance have been conducted under headphones using ''adjustment'' protocols in which listeners match the position of the fused image to that of a reference stimulus (or to midline) by varying a binaural parameter, such as interaural differences in time (ITD) or level (ILD), of either the lead or lag (von Bekesy, 1930; Wallach *et al.*, 1949; Haas, 1949; Snow, 1954b; Leakey and Cherry, 1957; Zurek, 1980; Yost and Soderquist, 1984; Shinn-Cunningham *et al.*, 1993, 1995). The dominance of the lead is quite compelling. For example, in their classic study, Wallach *et al.* (1949) reported combinations of lead and lag ITDs (each favoring a different ear) that resulted in a fused auditory image at the center of the head. At a delay of 2 ms an ITD of 100 μs in the lead required an ITD of approximately 400 μs in the lag. The fact that the lag did succeed in centering the fused image suggests that the lead did not dominate perception completely. In fact, a simple linear weighting function would imply that the directional cues contained in the lead were weighted four times more heavily than those of the lag.

A more precise estimate of localization dominance can be achieved by asking subjects to ''point'' to the perceived location of the auditory image using an acoustic pointer whose ITD can be manipulated by the listener (e.g., Zurek, 1980; Shinn-Cunningham *et al.*, 1993). The advantage of this protocol is that the perceptual weight of the lead and lag can be directly estimated from the data using very few parameters. Using a simple model, Shinn-Cunningham *et al.*

(1993) succeeded in calculating weighting factors for the lead and lag from several published studies and found that typical values ranged from 80%–90% and 10%–20% for the lead and lag, respectively. Hence, Wallach *et al.*'s (1949) original finding, that while the lead dominates localization the lag contribution is also important, has been confirmed over the years in several labs using various techniques.

2. Azimuthal-plane studies in free fields and rooms

While headphone studies were successful in providing a quantitative measure of the perceptual weights of the lead and lag, they certainly did not provide a ''realistic'' acoustic environment. Free-field studies, on the other hand, do provide more realistic everyday scenarios, although they have their own limitations. Early studies in free field using a two-loudspeaker system showed that the perceived location of the fused image is dominated by the lead speaker (e.g., Wallach *et al.*, 1949; Leakey and Cherry, 1957; Snow, 1954a; Haas, 1951; Leakey, 1959). Recall that these stimuli are further unrealistic in that the lag is not attenuated relative to the lead; this is unlike a true reflection, which can be dramatically attenuated relative to the source (e.g., Fig. 1). A classic example of this effect was demonstrated by Leakey and Cherry (1957) for loudspeakers located at 45° to the right and left, using speech signals. With no delay, the fused image appeared at a central location in front of the listener, and a delay of a few milliseconds in one loudspeaker shifted the entire auditory image to the other, leading speaker. Leakey and Cherry also showed that increasing the level of the delayed signal by several decibels shifts the fused image back to its central location. Similarly, Haas (1951) demonstrated that when the level of the leading source is decreased there appear to be two simultaneous sources in the directions of the two speakers. Hence, there is trade-off in localization dominance between delay and relative signal levels of the lead and lag.

To date, one published study has been conducted in the azimuthal plane with more than two speakers. Figure 4 shows data from Litovsky *et al.* (1997b) with three speakers (nine possible lead–lag combinations) at delays of 0 to 10 ms. In the azimuthal plane (open symbols) at delays of 1–2 ms, the leading source location was chosen on 95% of trials, providing strong evidence that the lead dominated localization. At longer delays (above 5 ms), both lead and lag locations were chosen equally, suggesting that localization dominance was no longer effective (listeners heard two sounds but could not determine which was the leading source). Results of free-field studies are generally consistent with the headphone studies; however, the techniques have been less sophisticated and have yielded little information regarding the relative weights of the lead and lag in localization.

Finally, a related auditory illusion known as the ''Franssen effect'' occurs with similar stimulus conditions in which the first sounds arriving at a listener dominate spatial perception (Franssen, 1963; see Hartmann and Rakerd, 1989; Yost *et al.*, 1997). This illusion refers to the finding that the location of a long duration tone at one spatial location is identified as arising from the location of a short tone burst that precedes the longer tone. For example, if a 50-ms tone at one

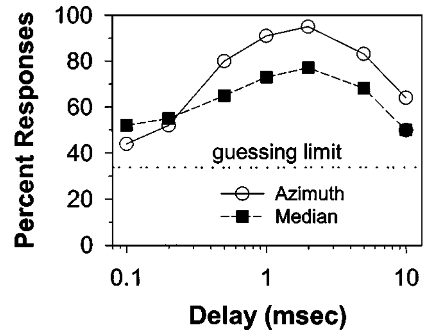

FIG. 4. Localization dominance measures in free field. Lead and lag locations varied in 90° steps along each plane (0, ±90° in azimuth and 0, 90, 180° in median). Listeners reported the location that was nearest to that of the fused event. Mean responses for eight subjects are plotted. For both source planes, the plots show the percentage of trials in which the leading source was nearest to the sound image (reprinted from Litovsky *et al.* 1997b, with permission).

location is presented with a sudden onset and a gradual offset, and at a different location a tone is turned on gradually and remains on for a long time, listeners will perceive a short tone at the onset from the leading speaker, then a room-filling diffuse ongoing sound. When forced to choose a location, subjects resort to the initial tone location. The illusion works for midfrequency tones presented in reverberant rooms. It does not work for noise stimuli, for stimuli presented in anechoic rooms or over headphones, or when the tones are low or very high in frequency (Hartmann and Rakerd, 1989; Yost *et al.*, 1997). The ability to produce the Franssen illusion is correlated with the ability of listeners to localize sounds, in that sounds that are difficult to localize appear to generate the strongest Franssen illusion (Hartmann and Rakerd, 1989; Yost *et al.*, 1997). Thus the Franssen effect is consistent with localization dominance in the precedence effect, in that the first information arriving at the listener controls the reported location of the perceived sound.

3. Free-field studies in the median sagittal plane

Since reflections in natural rooms arrive from multiple directions, including the walls, ceiling, and floor, they are likely to contain directional information that includes binaural cues (ITD and ILD) as well as monaural spectral and level cues. Until very recently, studies on localization dominance have focused exclusively on stimuli containing binaural cues, simulating reflections arriving from the walls. Little attention has been paid to scenarios in which reflections might contain few binaural cues and the primary directional cues are monaural spectral, such as when they arrive from the ceiling and floor directly in front of the listener. To the extent that models of localization dominance exist, they have thus focused on binaural mechanisms.

To date, a handful of studies have attempted to measure localization dominance under conditions in which monaural spectral cues are prominent. The first such study (Blauert, 1971) demonstrated that when the lead and lag are presented in the median sagittal plane from 0° (front) or 180° (back) and the lead–lag delay is greater than 0.5 ms, the lead dominated the perceived location of the fused image. However, the longest delay used in that study was 0.88 ms, which is within the ‘‘summing localization’’ window and not quite within the temporal region normally used in azimuthal/binaural studies on localization dominance (Blauert, 1997). Several recent studies (Rakerd and Hartmann, 1997; Litovsky *et al.*, 1997b; Dizon *et al.*, 1997) measured localization dominance in the median sagittal plane at delays greater than 1 ms. These studies employed a multiple-loudspeaker paradigm with various combinations of lead–lag locations, and listeners were instructed to report which of the loudspeakers was nearest to the perceived location of the fused sound image. The filled symbols in Fig. 4 show the results from Litovsky *et al.* (1997b) in the median-sagittal plane. At delays of 1–2 ms the leading source location was chosen on a majority of trials, as was found in the azimuthal plane. Thus the basic effect of localization dominance occurs at similar delays in the azimuthal or median planes. The effect does seem a bit weaker in the median plane, possibly since localization is poorer in the median plane, and adding a reflection creates coloration caused by a spectral ripple (Yost *et al.*, 1996) which can lead to localization errors (Blauert, 1997).

C. Lag discrimination suppression

While studies on fusion and localization dominance measure perceptual effects that are somewhat analogous to our experience in reflective environments, there exists a third approach to probing the auditory system’s sensitivity to stimulus parameters of the lead or lag. These experiments on lag discrimination suppression measure the ability of listeners to process spatial information about the lagging stimulus (relative to the leading stimulus). Experiments have been conducted both under headphones and in free field. Headphone experiments measure just-noticeable-difference in the ITD or ILD of the lead and lag (e.g., Zurek, 1980; Gaskell, 1983; Yost and Soderquist, 1984; Saberi and Perrott, 1990; Shinn-Cunningham *et al.*, 1993; Saberi, 1996; Tollin and Henning, 1998), whereas free-field experiments measure discrimination of positional changes (e.g., Perrott *et al.*, 1989; Freyman *et al.*, 1991; Litovsky and Macmillan, 1994; Yang and Grantham, 1997a, 1997b; Litovsky, 1999). In general, when the delay is short (less than 5 ms for clicks), changes in the lag are extremely difficult to discriminate; changes in the lead are much easier to discriminate, although performance is somewhat worse than on single-source discrimination. As Litovsky and colleagues have pointed out (Litovsky and Macmillan, 1994; Litovsky, 1997; Litovsky and Ashmead, 1997), while the lead-discrimination task requires listeners to ignore irrelevant information in the lag, the lag-discrimination task requires that listeners extract information from a sound that is not audible as a separate auditory event. Based on what is known about localization dominance, a prediction can be made that lead discrimination should be somewhat hampered by the presence of the lag, since the latter is not completely suppressed. In addition, lag discrimination should be the worse condition since directional information contained in the lag is the primary target of the suppressive mechanism.

1. Studies using headphones

The advantage of headphone studies is that they allow careful stimulus control and independent manipulation of each directional cue. The classic demonstration of the discrimination effect was Zurek’s (1980) report that listeners’ ability to discriminate changes in ITDs or ILDs of the lagging source deteriorates in the presence of the lead. The effect is largest at delays of 2–3 ms, the same delay range at which localization dominance and fusion are also known to be quite strong (see Secs. I A and I B).

Various aspects of this finding have since been replicated, with some caveats regarding the exact delays at which the effect is strongest (e.g., Gaskell, 1983; Tollin and Henning, 1998), limitations of the procedure due to high inter- and intra-subject variability (Yost and Soderquist, 1984), and possible effects of training (Saberi and Perrott, 1990). One of the intriguing outcomes of such studies has been a report that discrimination performance can be directly related to measures of localization dominance. Shinn-Cunningham *et al.* (1993) developed a simple model for localization ‘‘adjustment’’ data, which was able to accurately predict performance on the discrimination task, suggesting a tight relation between these two measurements. Thus when the leading source dominates the perceived location of the fused image, listeners are also unable to extract directional cues from the lag as easily as they can from the lead. This is consistent with Zurek’s (1980) postulation that the precedence effect results from a temporary loss of sensitivity to interaural cues shortly after the onset of the lead.

2. Studies in free field

Discrimination studies in free field can be categorized into two complementary experimental paradigms. One approach, which is analogous to ITD or ILD just-noticable-differences (JNDs), has been to measure the minimum audible angle (MAA; estimating the smallest change in the direction of a sound source that can be reliably detected) for the lead, lag, or a single source sound. In these studies the delay is kept short (2–4 ms), within the range of delays where fusion is known to be strong. Numerous studies have shown that lag-discrimination MAAs are higher than single-source MAAs by a few degrees (Perrott *et al.*, 1987; Perrott and Pacheco, 1989; Perrott *et al.*, 1989; Litovsky and Macmillan, 1994; Litovsky, 1997). In addition, lead-discrimination MAAs are better than lag-discrimination, but still somewhat worse than single-source MAAs (Litovsky and Macmillan, 1994; Litovsky, 1997). These results suggest that, while the presence of the lead renders directional information present in the lag difficult to access, the lag also interferes with lead discrimination. Hence, the precedence effect does not represent complete suppression of the lag, but rather a strong dominance by the lead. A useful method for normalizing performance is the ratio of lead/single or lag/single thresholds (Litovsky, 1997; Tollin and Henning, 1998), which is appropriate for comparing performance across conditions or between groups of various ages or with differing amounts of hearing loss.

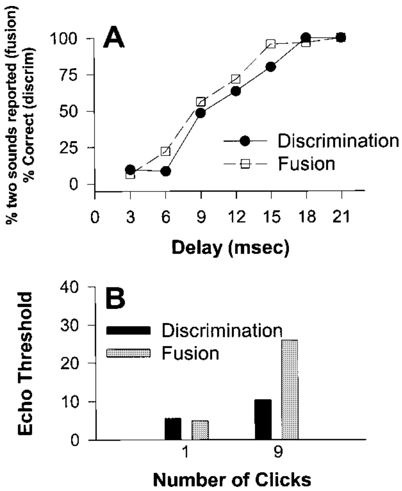

FIG. 5. Discrimination suppression measured in free field is similar to fusion under some conditions, but not others. A: Listeners report how many sources they hear (fusion; squares) or discriminate between two possible lag locations (discrimination suppression; circles). Data were obtained using one pair of 4-ms noise bursts. Results are highly similar for the two precedence tasks (replotted from Freyman *et al.*, 1991, with permission). B: Echo thresholds (delays at which discrimination performance reached 75% correct, or at which listeners reported hearing ''two sounds'' on 75% of trials) are shown for 1-click and 9-click conditions (Replotted from Yang and Grantham, 1997a, with permission).

A second experimental paradigm in free field has been to measure discrimination of a fixed change in the lag location while varying the delay. This method was originally developed by Freyman *et al.* (1991) in an attempt to link discrimination suppression to fusion. Figure 5(A) shows results from Freyman *et al.* (1991) which suggest that, under some conditions, echo threshold obtained in the fusion task is similar to the thresholds obtained with the discrimination task (defined as the delay for which d' equals 1). Freyman and colleagues postulated that, if listeners can subjectively ''hear out'' the lag as a separate sound then they should be able to extract directional information from the lag. Conversely, if the lag is fused with the lead then discrimination performance should be at chance. It turns out that this relation between fusion and discrimination holds true under specific conditions which depend on the locations of the stimuli, the number of stimuli presented within a given trial, and possibly other variations not studied.

There are two very interesting parameters that suggest a decoupling between auditory mechanisms mediating fusion and discrimination. The first parameter is the number of clicks presented within each trial. It appears that while the strength of fusion (measured with echo thresholds) increases substantially as the number of lead–lag pairs increases (further explored in Sec. II), discrimination suppression is less dependent on this stimulus variable [Yang and Grantham, 1997a; see Fig. 5(B)]. This effect relates to the influence of prior stimuli which are discussed in more detail in Sec. II.

A second situation in which there appears to be decoupling between fusion and discrimination is one in which stimuli are presented in the median-sagittal plane. As was discussed earlier, fusion is observed at nearly identical delays in the azimuthal and median planes, where binaural and monaural spectral cues, respectively, dominate localization.

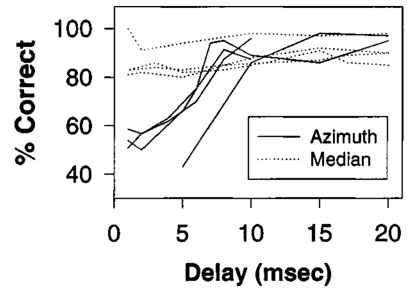

FIG. 6. Discrimination suppression in the azimuthal (solid) and median sagittal (dotted) planes for four listeners. Percent correct is plotted as a function of lead–lag delay (from Litovsky and Colburn, 1998; after Litovsky *et al.*, 1999).

In contrast, discrimination suppression has not been found in the median plane. Figure 6 shows results from a recent study by Litovsky *et al.* (1999), who compared discrimination performance in the two planes. Consistent with previous findings, in the azimuthal plane performance is poor at short delays and improves dramatically by 5–10 ms (e.g., Freyman *et al.*, 1991; Clifton *et al.*, 1994; Yang and Grantham, 1997a). In contrast, in the median-sagittal plane, performance is roughly independent of delay. When asked to describe which auditory cues were used during the task, all subjects reported that in the azimuth the only reliable cue was lateral movement of the sound, but in the median plane perceptual changes in the pitch of the fused auditory event were quite obvious and reliable. The authors postulated that in the median plane listeners were able to rely on cues provided by directional filtering by the pinnae, which are known to be important for identifying the elevation of sources (Fisher and Freedman, 1968; Blauert, 1969; Butler, 1969; Shaw, 1974; Gardner and Gardner, 1973; Middlebrooks and Green, 1991). Given the nature of the objective lag-discrimination task, where feedback is provided and listeners can use whatever cues are available to them, the delay dependence may only be robust when binaural cues are being varied, such as in the azimuthal plane in free field or using ITDs and ILDs under headphones. In fact, a recent report (Freyman *et al.*, 1998) suggests that also in azimuth, listeners are quite sensitive to various aspects of the lag, including its intensity and spectral content. Taken together, these findings are consistent with the notion that although we are usually not actively aware of reflections, we remain sensitive to information carried by these sounds. This sensitivity can be useful for enhancing certain information carried by the source and for a person's awareness of room acoustics.

D. Cross-frequency effects and uncorrelated lead–lag

Most studies on precedence have used idealized short-duration stimuli such as clicks. However, it is clear that in order to understand how reflections are processed by the auditory system, more realistic stimuli should be used. For instance, one could ask whether the reflection must be an identical copy of the source, which can be accomplished by using stimuli with different spectral contents or that are noncorrelated. Several such studies have been conducted, and taken together, they suggest that all aspects of the precedence effect are strongest when the lead and lag are identical, but to some extent they also operate when the lag is not an exact replica of the lead. In fact, it would be strange if precedence failed completely unless the lead and lag were identical. In most reverberant rooms the acoustics are such that some frequencies contained in the source are reflected while others are absorbed; hence, reflections are rarely identical to the source. However, realistic reflections can only be comprised of frequencies originally contained in the source, thus one might expect noncorrelated lead and lag signals to yield weak precedence.

To our knowledge, only one study has been conducted on cross-frequency effects using the fusion paradigm. Perrott *et al.* (1987) compared performance in free field for broadband noise bursts (50-ms duration) that were correlated to various gradations, and showed that fusion is somewhat weaker when the lead and lag consist of uncorrelated tokens of noise than when the tokens are correlated. In the uncorrelated condition, listeners reported hearing two sounds on most trials, regardless of delay; at delays below 10 ms there was a small proportion of trials on which one sound was reported. In contrast, in the correlated condition listeners experienced the ''classic'' fusion phenomenon; at short delays they reported hearing one sound on the majority of trials, and at longer delays they reported two sounds on the majority of trials.

The effects of the relative frequency of lead and lag components have been explored in more detail using the localization dominance and discrimination suppression paradigms. Scharf (1974) measured localization for a pair of tones presented from 45° to the right and left in an anechoic room. When the tones were the same frequency, a single fused image was heard from the leading speaker. As the frequency difference increased, the effect weakened, but was still present for tones differing by 1900 Hz. Blauert and Divenyi (1988) used virtual sound stimuli to show that the amount of lag-discrimination suppression increases as the overlap between the lead and lag spectra was increased. Like the fusion study noted above, this result is consistent with the notion that naturally occurring reflections (that contain more spectral overlap with the source) are more likely to be suppressed than unrelated stimuli. Similarly, in a headphone study of ITD discrimination, Divenyi (1992) found that lag-discrimination suppression is stronger if the lead frequency is lower than the lag. Divenyi fixed the lag frequency at 2000 Hz, varied the lead frequency from 500 to 3000 Hz, and found that sensitivity to lag ITD was poorest for the lowest lead frequencies. In fact, when the lead frequency was 2000 Hz or greater, there was little effect on performance, with ITD discrimination thresholds less than 100 μs. In contrast, lead frequencies of 500 or 1000 Hz resulted in lag ITD discrimination thresholds as high as 400 μs. This finding is somewhat unexpected if one assumes that precedence should be strongest for lead–lag stimuli that are similar in spectral content. However, Divenyi interprets his findings to suggest that discrimination suppression depends on the relative localizability of the lead and lag, rather than their spectral similarity. The lower-frequency lead stimuli are more easily localized than the higher-frequency stimuli, hence they produce stronger suppression of the directional information contained in the lag.

Similar results have been reported by others. Yang and Grantham (1997b) studied discrimination suppression for lead and lag stimuli that differed in their spectral overlap in both the free field (in an anechoic chamber) and with headphones (by manipulating interaural time differences). They showed that the parameters affecting discrimination suppression appear to be different in the two listening environments. In the free field, the amount of spectral overlap between the lead and lag stimuli was most effective in determining the amount of discrimination suppression as was found by Blauert and Divenyi (1988). However, over headphones, low-frequency information seemed to dominate high-frequency information even when there was little or no spectral overlap between the lead and lag stimuli. This dominance of low frequencies over high frequencies in lag-discrimination suppression is consistent with the fact that thresholds for processing interaural time differences are lower for low frequencies compared with those for high frequencies.

Using the pointer adjustment method with narrow-band noise bursts, Shinn-Cunningham *et al.* (1995) measured the extent to which the ITD in the lag influenced the perceived ITD of the fused image, and found that a high-frequency lag (1250 Hz) received almost no perceptual weight in the presence of a low-frequency (450 Hz) lead, but a low-frequency lag was perceptually weighted equal to a high-frequency lead. To test Divenyi's hypothesis that these findings arise from differences in the ''localization strength'' of components with different frequencies, Shinn-Cunningham *et al.* (1995) balanced the low- and high-frequency components by centering the image of simultaneously presented components with adjusted levels. The components were equally effective in centering an image when the level of the high-frequency component was greater than the level of the low-frequency component. When the levels of the lead–lag stimuli were adjusted in this way the asymmetry in discrimination was eliminated and the weight of the lead was the same for both cases.

The emphasis on the lead stimulus seen in localization results can also be seen in results related to the interaural cross correlation (Aoki and Houtgast, 1992). In measurements of the width or the compactness/diffuseness of the sound image, this study demonstrated that the dominance of the lead is the same for localization and diffuseness. These results suggest that the precedence effect, at least under some conditions, is not exclusively one of localization or lateralization *per se*, but rather a more general effect of reduction in the binaural information available in the lag.

These data suggest that suppression of the spatial information contained in the lag is strongest when the lead and lag are spectrally similar. However, there are still many unknowns about how the three measures of precedence depend on the spectral similarity between the lead and lag stimulus. Since real-world reflections are spectrally correlated, but not identical, additional information about precedence and spectral differences is necessary with the originating sound source in order to fully understand sound processing in reverberant environments.

E. Summary

We have reviewed psychophysical literature on perceptual phenomena which are based in single pairs of stimuli and which are thought to be related to the precedence effect. While fusion and localization dominance probe listeners' subjective impressions of how many auditory events are heard and what their perceived positions are, lag-discrimination suppression probe the extent to which aspects of the lag or lead are difficult to access due to the presence of the other. The discrimination studies avoid a problem inherent to fusion and localization dominance studies: the subjective nature of the tasks required of the listeners. The fusion and localization dominance tasks measure a perceptual ''impression,'' for which there is no objectively correct answer. Although subjective studies yield valuable information regarding auditory processing, the experimenter must be concerned with changes over time in listeners' criteria. On the other hand, discrimination experiments must be evaluated with full awareness that any of the multiple, changing aspects of the perceptions may be used as a cue. Thus performance may be related to timbre changes, for example, and not to spatial attributes directly.

Over the years, several attempts have been made to explain the time-course of precedence phenomena and to identify their locus in the auditory pathway. McFadden (1973) described a neural network that heavily weights the location of a sound source in the presence of reflections by inhibiting neurons that carry directional information regarding the opposite hemifield. Although this network does not take account of results in the median plane, it emphasizes the possible role of inhibition of directional information carried by the lag. The temporal window of precedence, which lasts for approximately 1–8 ms for clicks, was elegantly described by Harris *et al.* (1963) as a

> ''...neural gate previous to the place of binaural interaction that closes about 1 ms after the first neural response, permitting no further neural timing signals to be sent to the brain. This gate would reopen two, or slightly more, milliseconds later. A mechanism for such a neural gate could be self-inhibition and inhibition of lateral nerves by the nerves firing on the initial stimulus. This inhibition would have an inherent delay that would permit nerves to be fired in an interval of about 1 ms after the initial neural response. The inhibition would also last for about 2 ms.'' (p. 677)

This postulated inhibition might effectively reduce sensitivity to directional information, but not to other cues such as timbre, pitch, and loudness.

II. EFFECTS OF PREVIOUS STIMULATION

While most real-world sound sources and reflections occur only once or a few number of times, recent research on measures of precedence involving repeating the lead and lag stimuli several times have revealed some unexpected results.

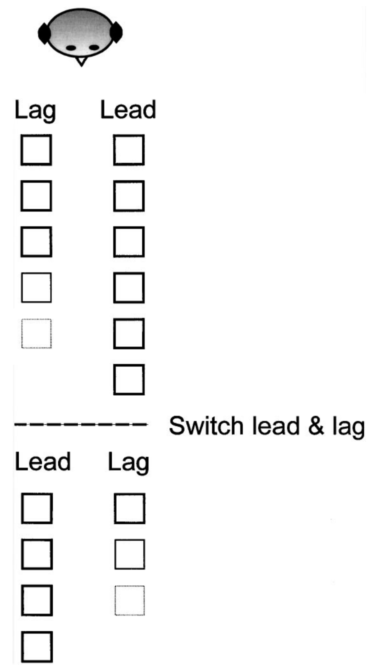

FIG. 7. This figure is a schematic cartoon describing the buildup and the breakdown of fusion. The first presentation of the lead and lag (at top) may not cause fusion leading to the perception of both stimuli. After several presentations fusion occurs. Following the switch of the spatial location of the lead and lag stimuli, fusion breaks down and is reestablished after the lead and lag stimuli are repeated at their new spatial locations.

These results reveal that the measures of precedence are dependent on listeners' immediate prior experience with the lead–lag presentations. If such experience is important for processing sound sources and their reflections then this implies that reflection processing may be more complicated than had previously been thought.

These recent data show that several measures of precedence change after several presentations of the lead–lag stimuli (buildup). For instance, after several repetitions of the lead and lag stimuli, listeners perceive a subjective ''fading out'' of the lagging stimulus (Clifton and Freyman, 1989). If, after the fade-out some parameters of the lead and lag are suddenly changed, such as their relative locations, then the lag reappears and is perceived again. Fusion is reestablished if the lead and lag stimuli are repeated several more times at their new spatial locations (Clifton, 1987). A schematic diagram of this phenomenon is displayed in Fig. 7. These findings suggest that aspects of the lag become more suppressed with number of presentations (buildup of precedence), and are no longer suppressed after certain parameters are switched (breakdown of precedence). These recent results are reviewed and the implications discussed below.

A. The buildup effect

Recent experiments have shown that, not only is echo threshold influenced by properties such as stimulus type, duration, and the spectral makeup of lead–lag stimuli, but short-term previous experience with a stimulus also influences echo threshold. In particular, it has been found that if a lead–lag click stimulus is repeated over and over again (referred to as a click train), echo threshold for the last lead–lag stimulus is raised by several milliseconds (Thurlow and Parks, 1961; Clifton and Freyman, 1989; Freyman *et al.*, 1991), with the most important parameter being the number of click-pairs in the train rather the train duration or the click rate (Freyman *et al.*, 1991). The slopes of echo threshold as a function of number of click-pairs in the train show the sharpest increase in threshold between 1–5 click-pairs, and an asymptotic value by 12 click-pairs, suggesting that echo threshold becomes stable following the acquisition of a certain amount of information about the lead–lag stimulus (Clifton and Freyman, 1997). This increase in echo threshold suggests that some sort of adaptation occurs over repetitions, such that listeners become less sensitive to reflections. In terms of fusion, lagging clicks are heard at first and then perceptually seem to fade away with stimulus repetitions. In terms of discrimination suppression, the time delay needed to discriminate the lag click increases as multiple repetitions of the stimulus occur. In terms of localization dominance, the image would be expected to move toward the leading source location, although measurements have not been made. This phenomenon has been called buildup of precedence.

Surprisingly, whether the lead stimulus originates from the left or the right side also affects the amount of buildup seen. This was first reported as a finding by Clifton and Freyman (1989), in which it was found that buildup was stronger when the leading click originated from the right side rather than from the left. Grantham (1996) pursued this notion and found that for a single noise burst no difference in discrimination suppression was seen between conditions in which the leading stimulus originated from the left or from the right side. If a judgment followed a train of lead–lag pairs (i.e., after buildup occurred), however, much more discrimination suppression was found on the right side than on the left side. Asymmetry in buildup has been suggested as evidence that more central brain mechanisms are involved in this phenomenon (Clifton and Freyman, 1989; Grantham, 1996).

Another stimulus characteristic which has been found to be important in buildup is the number of lagging clicks presented. Rather than presenting a simple lead–lag stimulus, Yost and Guzman (1996) presented multiple lags at different delays following the lead portion of the stimulus. Yost and Guzman found that, when the stimulus was a simple lead–lag click train (that is, only one lagging click was present), the time delay needed to see a buildup of fusion had to be much shorter than when two lag portions were present. For instance, no buildup of fusion occurred for a 12-ms separation between the lead and lag. However, if two lag clicks were presented, so that the first lag occurred 6 ms after the lead click and the second lag occurred 12 ms after the lead click (the same time separation as for the lead–lag stimulus), buildup was seen. This finding suggests that there is an interaction between the time separation and number of lagging clicks present.

Recent work suggests that the amount of buildup depends upon the type of task a listener is asked to perform (Yang and Grantham, 1997a). Although earlier studies have used both fusion and discrimination suppression tasks virtually interchangeably to measure buildup (Freyman *et al.*, 1991; Clifton *et al.*, 1994), Yang and Grantham suggest that fusion and discrimination suppression may be independent phenomena. They found that for a single lead–lag pair, performance on these two tasks was similar. But, with a train of bursts, fusion increased far more than discrimination suppression. This finding suggests that these two different aspects of precedence, fusion and discrimination suppression, may be governed at least to a certain extent by different mechanisms.

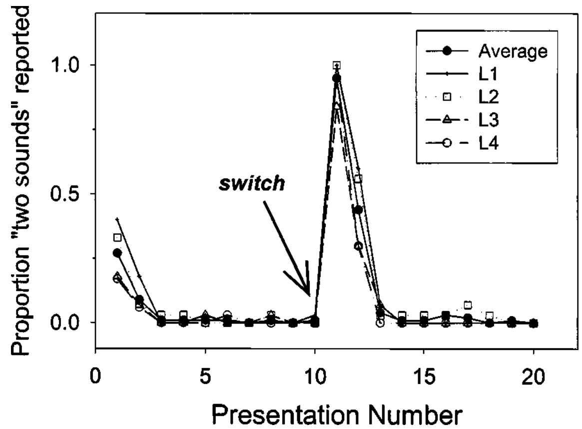

FIG. 8. The percent of times that ''more than one stimulus'' is reported as a function of the number of presentations of the lead and lag stimulus separated by 6 ms. Between the 10th and 11th presentation the locations of the lead and lag were reversed and remained reversed for the remainder of the presentations. Listeners indicated how many sound sources they perceived for the last lead–lag presentation for each repetition condition. When there are three or fewer presentations, listeners occasionally report perceiving more than one sound source. For additional lead–lag presentations only one sound source is reported (buildup). After the switch, most of the time more than one sound source is reported (breakdown of fusion). Then additional presentations yield reports of one sound source (buildup again). (Adapted from Yost and Guzman, 1996, with permission.)

Yang and Grantham (1997a; see Fig. 5B) have shown that when the lead–lag click-pair is preceded by a train of 9 click–pairs there is a strong dissociation between fusion and discrimination. Fusion echo thresholds increase by 15–20 ms. In contrast, discrimination echo thresholds are less affected by the preceding stimulus, and increase by 3–8 ms. It should also be pointed out that a change in the threshold for discriminating an interaural parameter or the spatial location of the lead and lag stimuli in these experiments does not necessarily mean that the lead or lag stimuli were heard as separate signals (i.e., that fusion did not occur). The discrimination can be based on the changes in the fused image such as the diffuseness or spatial extent of the perceived sound image, or lateral position of the fused event (see Yang and Grantham, 1997b).

B. The breakdown of precedence phenomena

It has been found that buildup of fusion or of discrimination suppression can be interrupted by presenting certain kinds of changes to the lead and lag portions of the stimulus, resulting in a dramatic decrease in echo threshold (cf., Fig. 7). This interruption of buildup is often called a ''breakdown'' or ''release from suppression.'' If the locations of the lead and lag suddenly switch, so that the lead now occurs where the lagging stimulus had been and vice versa (referred to as a reversal switch), fusion is disrupted and listeners hear both clicks. An example of this effect is shown in Fig. 8 with recent data from Yost and Guzman (1996). This release from suppression due to a switch in location was first reported by Clifton (1987) and has thus become known as the ''Clifton effect'' (Moore, 1996; Blauert, 1997). Other changes to the stimulus have also been shown to disrupt buildup. Clifton *et al.* (1994) found that if the time delay between the lead and lag of the last lead–lag pair was changed, discrimination suppression was reduced and listeners could detect the location of the lagging click more easily. Yost and Guzman (1996) found that only a particular kind of change in location results in a breakdown. If the lead and lag clicks are simply shifted one loudspeaker to the right (referred to as a lateral switch), there is no breakdown in fusion.

Hafter and his colleagues (see reviews in Hafter *et al.*, 1988; Hafter, 1997; Yost and Hafter, 1987) measured the ability of listeners to detect an interaural time difference (and in some experiments, an interaural level difference; see Hafter *et al.*, 1983) of a train of clicks. They consistently found that the early clicks in the click train contribute more than later clicks in the discrimination of a dichotic click train with an interaural difference than from a diotic click train without interaural differences. As one would expect, the more clicks there are in the click train (up to some limit), the lower the interaural difference thresholds become. However, for short temporal separations between the clicks in the train, the later clicks contribute less information than the earlier clicks. They developed a model which assumes the auditory system adapts to the early arriving interaural information, such that the early arriving binaural information dominates. While most of their work has been done using headphone delivered stimuli, Hafter *et al.* (1992) demonstrated that similar effects occur in an anechoic room. These results are like many cited in the precedence literature (see Hafter, 1997) in that spatial processing is dominated by early arriving information (the precedence effect), later samples have less binaural information than earlier samples, and there is a buildup of binaural information over repeated events.

Hafter and Buell (1990; see also Hafter *et al.*, 1988 and Hafter, 1997) studied stimulus conditions involving click trains and interrupted click trains. These experiments share many similarities to those measuring for the breakdown of precedence. For instance, suppose that a train of N clicks has an interaural time discrimination threshold of approximately 60 μs. If the train of clicks is made half as long, i.e., $N/2$ clicks long, the threshold can increase (for instance, to about 70 μs) since a short click train has less useful binaural information than a longer click train. However, if halfway through the long click train (after the $N/2$ click) noise is added, or something else (even silence) interrupts the train momentarily, the interaural-time difference threshold for the interrupted, long click train decreases to 50 μs. This indicates that the information in the interrupted click train is more effective than in the continuous train. It is as if the interruption caused the binaural system to restart its adaptation process and the click train is processed as two sets of $N/2$ clicks, rather than a long train of N clicks. This is similar to the Clifton effect, in which an interruption in the flow of repeated clicks causes a restart of the system that governs the way in which interaural information is associated with repeating click events.

C. Possible consequences of the buildup and the breakdown of precedence

Several ideas regarding processes involved in the precedence effect have emerged based upon findings of buildup and breakdown. The breakdown of suppression, especially the breakdown of fusion, suggests that the auditory system maintains information about the lag (reflection) stimulus, even when fusion occurs and localization is dominated by the lead stimulus. That is, for the early click events in a train of repeating clicks, the lag appears to be completely fused with the lead, and source location is dominated by that of the lead. However, when certain types of ''switches'' or inconsistencies involving the lead and lag take place, the lag is perceived as separate with its own perceived location. Thus it appears as if information about the lag is not eliminated, but rather that certain information about the lag is suppressed. Certain changes in the acoustic environment then release this suppression.

Phenomena such as left–right asymmetry and the relatively long time course over which buildup and breakdown occur suggest to some researchers that buildup and breakdown are governed by *central* brain processes (Clifton and Freyman, 1997; Grantham, 1996). It has been suggested that breakdown occurs with changes to the stimulus that are incompatible with a listener's immediate previous experience in an acoustic environment. Blauert and Col (1992) hypothesized that the precedence effect is at least partially controlled by processes in higher centers of the nervous system, and suggest that a model such as one involving a pattern recognition process may be appropriate to account for the listener's ability to selectively listen to the lagging source under certain conditions. Rakerd and Hartmann (1985) propose what they have termed the ''plausibility hypothesis.'' According to this theory, interaural parameters are assessed by listeners to determine their plausibility given the information listeners have about the environment (for instance, visual images). Thus interaural variables deemed to be implausible (such as a very large interaural difference of time) are discounted. Clifton and colleagues (e.g., Clifton *et al.*, 1994) suggest a similar hypothesis, in which reflections provide information about a listener's acoustic environment. Based on this (previous) information, listeners form ''expectations'' about the sounds that can occur given the acoustic environment. If a sound does not follow their expectations (such as in the case of a reversal switch), listeners reevaluate the acoustic stimulus, causing breakdown of fusion and/or of discrimination suppression. It should be noted that to date there is no evidence that either buildup or breakdown are learned effects or that they can be modified by practice (Clifton and Freyman, 1997).

The fact that buildup and the breakdown of precedence occur for several different conditions has been used to suggest that reflection processing relies on central mechanisms such as those that are involved with cognition (see Blauert, 1997). While this is one possible way to theorize about precedence mechanisms, data may be too sparse to resolve one theory over another. For instance, if buildup and breakdown of precedence cannot be modified by practice or other conditions that control learning, then many theories of cognition would not be applicable for dealing with these precedence effects. Precedence buildup and breakdown suggest that we do not fully understand the basic processes governing precedence, especially fusion. If fusion changes as a function of immediate prior experience with the lead and lag stimuli, what aspects of the lead and lag (or the environment in which the lead and lag occur) cause fusion? A full understanding of reflection processing in real-world situations requires knowledge of precisely which aspects of the lead and lag alter the effects of precedence.

III. ANIMAL BEHAVIOR AND HUMAN DEVELOPMENT

Early behavioral-ablation studies showed that unilateral ablation of the auditory cortex severely impairs the ability of cats to accurately choose the hemifield containing the leading source, while the ability to find the hemifield containing a single-source sound was not affected (Cranford *et al.*, 1971; Whitfield *et al.*, 1972). Performance was consistent with elimination of lead dominance in localization on the side of the lesion. The deficit was limited to cases in which the leading source was in the contralateral hemifield (opposite to the lesion site). For instance, if the left auditory cortex was ablated, left–right lead–lag pairs of tone pulses (23-ms delay) were localized at the leading source but right–left lead–lag stimuli were not localized consistently. Later studies (Cranford and Oberholtzer, 1976; Whitfield, 1978) found this effect in some but not all animals. These differences have been attributed to differences in the extent of the lesion, individual differences in behavioral learning strategies amongst cats, and differences in testing procedures.

Behavioral correlates of localization dominance and fusion have been found in several animals at delays that are similar to those reported for humans. Cranford (1982) made psychophysical measurements of localization dominance in normal animals. Click trains were presented with varying lead–lag delays (from 0.2 to 9 ms) either with a speaker on the left leading a speaker on the right, or vice versa. Cats were trained to release a foot pedal on the right if they heard a sound on the right, and another foot pedal on the left if they heard a sound on the left. Cats were only rewarded if they released a foot pedal corresponding to the side of the leading click. (For the case of 0-ms delay, cats were rewarded for releasing either foot. This symmetric case was run in separate sessions.) Mean results from six cats tested in the Cranford study and from studies in rats and humans are shown in Fig. 9. In the Cranford study (circles), all animals responded mostly to the leading side at lead–lag delays of 0.5 to 2 ms. At longer delays some cats still responded only to the lead, while others responded to both lead and lag. These results for cats are very similar to the human localization-dominance results (Litovsky *et al.*, 1997b) plotted in the same figure. It appears that the delays at which cats seem to show localization dominance are about the same as the delays for which human listeners experience localization dominance and fusion.

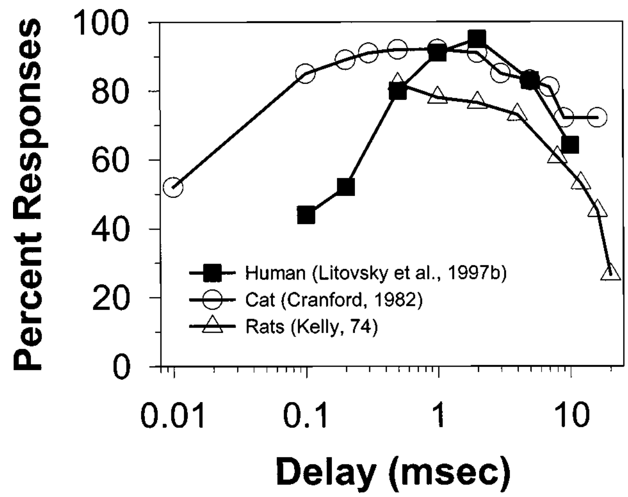

FIG. 9. Behavioral measurements in animals. Data for three species are compared: humans (square; from Litovsky *et al.*, 1997; replotted with permission), cats (circle; from Cranford, 1982; replotted with permission) and rats (triangle; from Kelly, 1974; replotted with permission). In each case performance is plotted as a function of lead–lag delay. Human data are azimuthal data replotted from Fig. 5; percent of trials in which listeners reported that the sound was located at the lead speaker are plotted. Cat data represent the percent of trials in which cats oriented toward the leading speaker. Rat data represent animals' performance (rated from 0%–100%) score for discrimination of right–left (lead–lag) trials from left–right trials.

Other studies reported similar findings in rats (Kelly, 1974, Fig. 9, triangles) and crickets (Wyttenbach and Hoy, 1993). Finally, Keller and Takahashi (1996) measured sound localization in owls for lead–lag stimuli. Since the owls were asleep or not appropriately positioned on 90% of trials, the behavioral data are scarce in seven of the ten animals tested. In three animals, at lead–lag delays of 2 to 10 ms the owls turned their heads consistently toward the leading speaker on the majority of trials. Most recently, Populin and Yin (1998) used a magnetic search coil technique to measure eye movements in cats whose heads were fixed in place. These cats were able to localize a single-source click or noise burst to within a few degrees. Lead–lag trials included two sources placed at 18° to the right and left. At a delay of 0 ms, cats oriented their eyes toward midline, suggesting that they heard a ''phantom'' source at that location. As the delays were increased, the eye movements shifted gradually toward the leading speaker, approximating single source by 300 μs.

Hochster and Kelly (1981), motivated by the earlier ablations studies in cats, measured localization dominance in children 6–16 years of age with temporal-lobe epilepsy. For stimuli consisting of single clicks, these children were able to identify which of two loudspeakers contained the stimulus. On precedence trials containing lead–lag stimulus pairs with short delays (1–4 ms), both epileptic and normal children were able to identify the leading speaker on nearly 100% of trials. However, the effect was substantially weaker in the epileptic children than normal children or adults, suggesting that the temporal lobe might be functionally involved in mediating localization dominance.

A series of precedence studies in human infants and children and in newborn and young dogs were undertaken, partly motivated by the animal behavior studies (for extensive reviews see Clifton, 1985; Litovsky and Ashmead, 1997). Localization dominance, measured in much the same way as it was in the cats, does not seem to be present in newborn human infants (Clifton *et al.*, 1981), although single-source discrimination occurs within hours after birth (see Clifton, 1985). Localization dominance is first measured in humans at 4 to 5 months of age, but at that age fusion echo thresholds are quite high (25 to 45 ms) compared with adult thresholds measured under the same conditions (8 to 15 ms). Echo thresholds in human infants were measured using a conditioned head-turning task. Lead–lag sounds were presented from loudspeakers located at 90° to the right and left and their relative onsets were delayed by 5–50 ms. At the onset of each trial there was a 7-ms delay, which had previously been shown to produce effective fusion in infants. The delay was abruptly lengthened; when infants heard the lagging sound at its location they responded by turning their heads toward this ''novel'' sound. Echo thresholds in 5-month-old infants are approximately 26 ms for click stimuli, which is significantly higher than adults' thresholds of 9 ms. By 5 years of age, adult thresholds are reached for click stimuli, however, for long duration more complex stimuli echo thresholds are significantly higher in children than in adults (Morrongiello *et al.*, 1984). Behavioral studies on localization dominance have also been conducted in young dogs, who do not show any evidence of precedence through the fifth month of life (Ashmead *et al.*, 1986).

Discrimination suppression also undergoes significant developmental changes during early childhood. Litovsky (1997) measured MAAs in free field for the leading source (in the presence of the lag at midline), for the lagging source (in the presence of the lead at midline), and for a single source, at ages 18 months, 5 years, and adult. The stimuli were 25-ms noise bursts and the lead–lag delay was 5 ms. Since this lead–lag delay is below the echo threshold for these stimuli, the image is fused. Single-source MAAs are adultlike (1 to 2°) at 5 years of age and fairly low (5°) by 18 months of age. Lead MAAs are quite low in adults (1.7°), somewhat elevated at 5 years (4.4°), and substantially higher at 18 months (23°). Lag MAAs are still low in adults (1.7°), but substantially higher in 5-year olds (27.5°) and 18-month olds (65°). Lead MAAs reflect listeners' ability to focus on the first-arriving wave front and to discriminate between leading source locations in the presence of the lagging source, and this ability improves dramatically with age. Lag MAAs reflect listeners' ability to extract directional information from a sound that is not heard as a separate auditory event; this ability improves somewhat with age but is still quite underdeveloped at 5 years relative to adults. These findings are consistent with the fusion echo threshold data which were obtained using a task that measures children's ability to localize the lag as a separate sound. In the MAA study, normalizing the lead and lag results by the single-source results maintained the developmental differences observed, suggesting that lead and lag MAAs are not merely the ''by product'' of a ''noisy'' single-source discrimination ability. As Litovsky (1997) points out, while the developmental work may point to maturational changes in the central auditory pathway, attentional and learning processes cannot be ruled out.

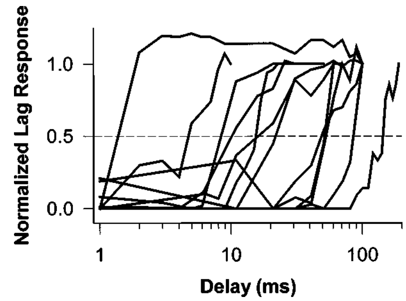

FIG. 10. Examples of recovery functions of neural half-maximal ISDs are plotted. For all neurons, both leading and lagging stimuli were at locations or ITDs which elicited robust responses with single stimuli. The recovery functions represent the normalized lagging responses as a function of ISD. (Replotted from Litovsky and Yin, 1998a, with permission.)

IV. PHYSIOLOGICAL CORRELATES OF PRECEDENCE

In recent years there has been a surge of activity exploring physiological substrates of precedence. Two decades ago (Altman, 1975), and more recently (Carney and Yin, 1989), there were suggestions made that such a substrate might be found in responses of single neurons in the central nucleus of the inferior colliculus (IC). Using stimuli that consisted of a single click to each ear with unusually long ITDs (tens of milliseconds), Carney and Yin (1989) found that the response to the lagging click was suppressed in most cells, even when the leading click did not elicit a response. In addition, at very short delays (less than 1 ms), the lagging click often produced a backward suppression of the leading click. Carney and Yin (1989) hypothesized that while the long-lasting forward suppression might reflect a neural correlate of echo suppression, the backward suppression may be important for summing localization.

Several years later, Yin (1994) tested that hypothesis by presenting stimuli from two locations (or with different ITDs) during the time period associated with summing localization (less than 1 ms), when listeners hear a fused auditory event at a phantom location near the leading speaker. Several IC neurons gave responses consistent with summing localization: as the lead–lag delay increased from zero, the responses to the lead–lag combination followed the response of the neuron to ITDs in between the lead and lag delays, progressing with increasing delay to the response expected from the lead alone. Yin (1994) suggested that the discharge rates of these neurons might be related to the location at which the animal would perceive the phantom sound source. The same study also explored neural responses to lead–lag stimuli at longer delays. Note that IC neurons are usually sensitive to specific ITDs and azimuthal locations, hence measurements were consistently made for lead and lag locations at each neuron's ''best'' azimuth or ITD. At short delays most neurons responded only to the leading source, and as the delays were increased the lagging response recovered from suppression, resulting in recovery curves much like the fusion psychometric functions seen in human listeners. For the 65 cells studied, the half-maximal delays (at which the lagging response recovered to half of the nonsuppressed response) ranged from 1 to 100 ms with a median of 20 ms.

Similar results were reported by Litovsky and Yin (1998a) for a population of 94 neurons studied in free field [e.g., Fig. 3(A)] and with dichotic stimuli [e.g., Fig. 3(B)]; half-maximal delays ranged from 1.5 to 154 ms with a median of 27 ms. Figure 10 shows normalized lagging responses (lag response at each delay divided by lag response at maximal delay) from a sample of 12 neurons in the IC. The data suggest that all neurons show strong suppression at the shorter delays tested and no suppression at long delays. In addition, there is substantial variability in delays at which IC neurons exhibit suppression of the lagging response. (In fact, this variability is also apparent in Fig. 12, where IC data are compared with data collected at other levels in the auditory pathway.)

Seeking further physiological correlates of precedence, Litovsky and Yin (1998a, 1998b) measured responses of IC neurons using stimulus parameters that are known to influence echo thresholds in human listeners. They reported a tendency for suppression to last longer at lower overall stimulus levels, which is consistent with psychophysical results for localization dominance (e.g., Shinn-Cunningham *et al.*, 1993). Suppression was also stronger for noise bursts than for clicks, for long-duration noise than for short-duration noise, and when the leading level was increased, all of which are consistent with psychophysics (see Blauert, 1997, and Table I). The most striking correlate was obtained for stimuli presented either in the azimuthal plane or in the median-sagittal plane (see also Litovsky *et al.*, 1997b). Comparisons were made by positioning the lagging source at 0 azimuth–0° elevation (a location common to both planes), and the leading source at locations along the azimuth and median planes that produced similar discharge rates. For the 39 neurons studied there was a high correlation in half-maximal delays ($r=0.8$) for the two planes. These results are consistent with the psychophysical measurements of fusion (Litovsky and Colburn, 1998) and localization dominance (Litovsky *et al.*, 1997b; see Fig. 4) in the azimuthal and median planes. The combined physiological and psychophysical findings suggest that fusion is mediated by the same neural mechanisms regardless of whether binaural disparity cues (azimuth) or spectral cues (elevation) are prominent. As was discussed above (Sec. I B 3), Litovsky *et al.* (1997b) have argued that models of fusion which assume that interaural delays are an integral aspect of the precedence effect (Lindemann, 1986a; Franssen, 1963; Shinn-Cunningham *et al.*, 1993) do not address the similarity of fusion for stimuli in vertical and horizontal planes. Models which integrate spectral cues are yet to be developed.

The studies discussed thus far were conducted in barbiturate-anesthetized animals. The potential problems in data interpretation stems from the fact that barbiturates are thought to enhance physiological inhibition in the central auditory pathway (Kuwada *et al.*, 1989). Responses of IC neurons to lead–lag stimulus pairs have also been measured in awake rabbits (Fitzpatrick *et al.*, 1995) and owls (Keller and Takahashi, 1996). In the awake rabbit, half-maximal delays in 55 neurons (using dichotic click stimuli with both lead and lag at the neuron's ''best'' ITD) ranged from 1 to 64 ms with a median of 6.3 ms, which is somewhat lower than the values obtained in anesthetized cats. In the awake barn owl (Tyto alba), Keller and Takahashi (1996) measured responses of 51 neurons in the external nucleus of the inferior colliculus (ICx) to pairs of noise bursts (3 or 100 ms) with lead–lag delays of 0.5 to 5 ms. These neurons, which are excited maximally by stimuli presented from specific locations in space, responded only to the leading stimulus at these short delays. While some neurons' lagging response recovered by 5 ms, others did not. Behaviorally, at these same delays the owls only turned their heads toward the leading source (see Sec. III above).

As several authors point out (e.g., Fitzpatrick *et al.*, 1995; Litovsky and Yin, 1998a), it is difficult to determine whether the differences stem from species differences or the effect of barbiturates. One of the challenges in trying to relate neural evidence of echo suppression to behavioral echo thresholds, is that for brief stimuli a large proportion of neurons recover from suppression at delays that extend far beyond behavioral echo thresholds. Behavioral echo thresholds for brief stimuli are usually reached by 5 or 10 ms in humans and other animals. However, the proportions of neurons with half-maximal delays less than 10 ms ranged from approximately 15% (Yin, 1994) to 22% (Litovsky and Yin, 1998a) in anesthetized cats, and reached over 30% in awake rabbits (Fitzpatrick *et al.*, 1995). It has been argued (Yin, 1994) that perceptual echo thresholds are most likely generated by those neurons with the lowest half-maximum delays in much the same way that behavioral thresholds of pure tones have been related to physiological thresholds of auditory nerve fibers (Liberman, 1978). It has also been noted that several aspects of precedence, such as accurate localization of the lag at its respective position and equal-loudness perception of the lead and lag, are not released from the influence of the lead until the delays are in the tens of milliseconds, which may account for the neurons with long half-maximal delays (Fitzpatrick *et al.*, 1995; Litovsky and Yin, 1998b).

The extent of suppression in the IC also depends on the location or ITD of the leading stimulus, but the relationship between the lead and lag locations seems to affect neurons in a variety of ways. In a headphone study using dichotic clicks, Fitzpatrick *et al.* (1995) reported that while half of the neurons show longer suppression when the ITD is near the ''best'' ITD, the other half show stronger suppression with the lead near the ''worst'' ITD. The ''worse'' neurons, which never fire in response to the lead, provide the strongest evidence that the suppression observed in ICC neurons is not merely a function of refractoriness, whereby once a neuron has already fired in response to the leading source the probability that it will fire in response to the lagging source is diminished. Litovsky and Yin (1998b) made measurements both dichotically and in free field. In the latter case the lag was held constant at each neuron's ''best'' location and the lead location was varied along the azimuth in 15° increments. When considering only the neurons whose lead response is modulated with azimuth, they found that the great majority (85%) showed maximal suppression when the lead was at the neuron's ''best'' location, and a minority showed maximal suppression when the lead was at the neuron's

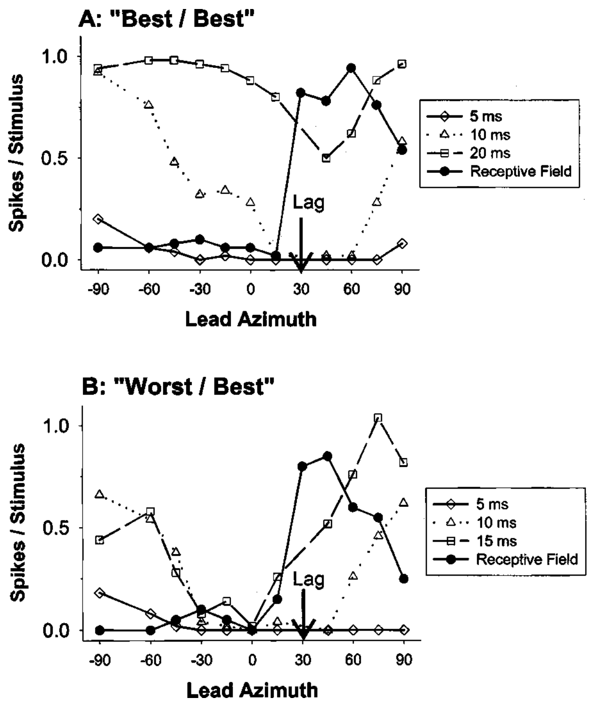

FIG. 11. Examples of neurons with ''best'' and ''worse'' lead location suppression. A: Example of a ''best/best'' neuron where suppression is strongest when the lead is in the neuron's ''best'' azimuthal locations (CF=3 kHz). Responses to single clicks are shown in dark circles, representing the neuron's receptive field. The arrow points to the location of the lagging source (+30°); if no suppression occurs then the lagging response should equal the response of the neuron to a single click at 30°. Responses to the lagging clicks at ISDs of 5, 10, and 20 ms are shown. At 20 ms, suppression is relatively weak and occurs only when the lead is at the best response area. At 10 ms, the suppression spreads out to the neuron's nonresponsive area, and at 5 ms suppression is almost complete, regardless of the lead location. B: Example of a ''worse/best'' neuron where suppression is strongest when the lead is in the neuron's ''worse'' azimuthal locations (CF=1.6 kHz). Suppression for this neuron also increases as the delays are decreased, with almost complete suppression at 5 ms, regardless of the lead location. (Reprinted from Litovsky and Yin, 1998b, with permission.)

''worse'' location. Examples of these two neuron types are shown in Fig. 11.

Of course, the full story is only told when all of the neurons in the population are considered, including those that do not show a directionally sensitive response to the lead. In a recent study using ''virtual space'' stimuli, Litovsky *et al.* (1999) quantified both the amount of response modulation (directional sensitivity) related to the excitation produced by the lead, and the modulation related to the suppression of the lag. In their population of IC neurons, 80% show modulation of both lead excitation and lag suppression, and 20% show modulation of either the lead excitation or the lag suppression. In their study, the directional properties of the lead stimulus were then digitally manipulated such that, as the azimuth of the lead was varied the ILD was held constant while the ITD and spectrum varied naturally. This manipulation resulted in loss of either lead excitation or lag suppression in the majority of neurons (83%). The authors have thus argued that the inputs mediating directional sensitivity to the source (excitation produced by lead) and the reflection (suppression of the lag) may be decoupled, and arrive from different populations of neurons.

Although the physiological recordings discussed above were made in the IC, which is a major site of binaural convergence and interaction in the auditory pathway, the generation of inhibitory effects may not necessarily be in the IC itself. Inputs to the IC arrive in the form of direct monaural input from the cochlear nuclei, indirect binaural input via the superior olive, and multisynaptic inputs via the lateral lemniscus. Auditory studies in the monaural pathway have shown that in response to click pairs with varying lead–lag delays suppression of the lagging response occurs in the auditory nerve (Parham *et al.*, 1996) and ventral cochlear nucleus (Wickesberg, 1996; Parham *et al.*, 1998). In these studies, the half-maximal suppression occurs at around 2–4 ms with full recovery seen for all fibers by 8 ms. In a slice preparation, Wickesberg and Oertel (1990) found putative inhibitory projections from one subdivision of the cochlear nucleus to another, which is most effective at interstimulus intervals of 2 ms; they have thus suggested that intrinsic inhibitory circuits in the cochlear nucleus might mediate a monaural correlate of echo suppression. Several other studies have shown that a neural correlate of forward-masking in response to tones can be seen in the cochlear nucleus (Boettcher *et al.*, 1990; Kaltenbach *et al.*, 1993; Shore, 1995). In the superior olivary complex (SOC), Fitzpatrick *et al.* (1995) found that most neurons recovered by 4 to 8 ms (median 1.9 ms) with monaural stimulation. However, the short time courses of monaural suppression compared with the much longer time course of suppression observed in the IC suggests that much of the suppression is generated at levels above the monaural pathway. In fact, reports of measurements made in the auditory cortex (Reale *et al.*, 1995; Fitzpatrick *et al.*, 1995) suggest that suppression can last for hundreds of milliseconds.

In an attempt to compare the amount of suppression that can be measured at various levels in the auditory pathway, we have replotted data from a number of studies conducted in the following areas: auditory nerve (derived from Parham *et al.*, 1996), antero-ventral cochlear nucleus (derived from Parham *et al.*, 1998), inferior colliculus (replotted from Litovsky and Yin, 1998a) and auditory cortex (from Reale *et al.*, 1995). Figure 12 shows population histograms of half-maximal delays at each of these stages in the auditory pathway. These panels point to the marked increase in suppression as one ascends the auditory pathway, going from monaural circuits to binaural ones. While suppression in the cochlear nucleus is no longer present by 10 ms, in the IC suppression is quite strong at those delays for most neurons, and cortical neurons do not begin to recover before 50 ms. Finally, given the extensive physiological measurements conducted to date, in Table II we summarize half-maximal delays observed at various levels in the auditory pathway, in a number of different species, and with either anesthetized or awake preparations.

Yin (1994) makes a convincing argument that long-lasting suppression observed in the IC is not due to long refractory periods caused by intrinsic neural mechanisms, and is therefore probably due to synaptic inhibition. The IC contains abundant inhibitory synapses, primarily from the lateral superior olive (LSO, e.g., Saintmarie, 1989) and the

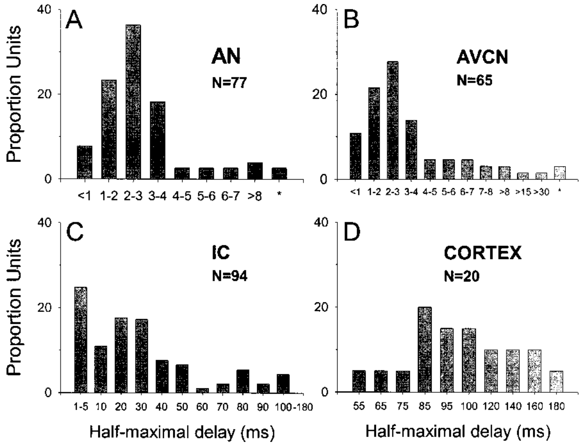

FIG. 12. Population histograms of half-maximal delays for click stimuli measured in cats at various levels in the auditory pathway. A and B: Recordings made in the auditory nerve (AN) and anteroventral cochlear nucleus (AVCN), respectively. In both populations stimulation was monaural and click stimulus level was 85 dB SPL peak *re* 20 μPa. AN data were derived from 77 fibers (with permission from Parham *et al.*, 1996; for two fibers values were not obtainable [*]). AVCN data were derived from 65 units (with permission from Parham *et al.*, 1998; for two units values were not obtainable [*]). C: Recordings made in 94 units in the inferior colliculus (IC) by Litovsky and Yin (1998a). Recordings were made using binaural stimulation with ITDs or azimuthal locations of both lead and lag placed in each neuron's maximal response area. Levels ranged from 50–80 dB. Data are reprinted with permission. D: Recordings made in 20 units of the auditory cortex by Reale *et al.* (1995). As in the IC, ITDs of both lead and lag elicited maximal responses. Data provided to the authors by Dr. John Brugge via personal communication, 1999.

dorsal nucleus of the lateral lemniscus (DNLL) (e.g., Adams and Mugnaini, 1984; Moore and Moore, 1987), as well as some that originate within the IC itself (Oliver *et al.*, 1994). Fitzpatrick *et al.* (1995) have postulated that the DNLL projections may play a prominent role in the results seen in the IC. Most neurons in the DNLL are sensitive to ITDs, hence the ''best'' ITD and ''worse'' ITD neurons could be invoked by ipsilateral and contralateral activation, respectively. Based on similar assumptions, Cai *et al.* (1998a) have successfully modeled both ''best'' ITD and ''worse'' ITD neurons, as well as other effects observed by Litovsky and Yin (1998a, 1998b), assuming ITD-sensitive inhibition from the DNLL.

To date, little else is known about the actual physiological mechanisms that mediate precedence phenomena. It is unlikely that perception can be explained solely on the basis of single-neuron responses, hence more sophisticated analyses such as determination of population codes may be necessary. Further, although initial stages of echo suppression may occur in the brainstem, single-neuron results cannot account for all perceptual phenomena related to precedence. For example, Litovsky and Yin (1998a) conducted an analysis which suggests that single neurons do not show signs of buildup. Perhaps the most convincing evidence that the IC may not account for behavioral measurements of precedence comes from a recent study by Litovsky (1998), who showed that neurons in the ICC of newborn cats exhibit the same type of suppressive effects as the adult neurons. This occurs as early as 8–9 days of age, before the full maturation of the structure or function of the auditory system and before the time that a young cat functions behaviorally in its auditory environment. Recall that human infants and young dogs do not exhibit localization dominance or fusion early in life, and that the phenomenon is only observed at 4–5 months of age. Thus, Litovsky (1998) concludes that behavioral manifestation of precedence is mediated at higher levels in the auditory pathway than the IC, which is consistent with the cat lesion studies and the human development studies, all of which suggest that the auditory cortex is essential for precedence to occur behaviorally.

V. MODELS AND THEORIES RELATED TO PRECEDENCE EFFECTS

There have been a few published models addressed to the precedence effect; however, there is no model currently published that is able to accommodate available data satis-

TABLE II. Physiological studies in mammals.

Recording site	Animal	State	Intra/ xtra	Stim presentation (Clicks)	50% recovery range (median)	Study
Auditory nerve	cat	decerebrate	xtra-cell	headphones	>1 to 20 ms (2.4)	Parham *et al.* (1996)
Cochlear nucleus (DCN)	cat	decerebrate	xtra-cell	headphones	>1 to 32 ms (2.7)	Parham *et al.* (1996)
Cochlear nucleus (VCN)	chinchilla	anesthetized	xtra-cell	headphones	1–2 ms	Wickesberg (1996)
Cochlear nucleus (PVCN)	mouse	slice	intra-cell	shock	2 ms	Wickesberg & Oertel (1990)
Superior olivary complex	rabbit	awake	xtra-cell	headphones	4–8 ms (1.9)	Fitzpatrick *et al.* (1995)
Inferior colliculus	cat	anesthetized	xtra-cell	headphones	1.5–100 ms (20)	Yin (1994)
Inferior colliculus	cat	anesthetized	xtra-cell	free field	1–90 ms (20)	Yin (1994)
Inferior colliculus	cat	anesthetized	xtra-cell	free-field azimuth	2–183 ms (28)	Litovsky & Yin (1988a, 1998b)
Inferior colliculus	cat	anesthetized	xtra-cell	free-field elevation	1–76 (16.5)	Litovsky & Yin (1998a, 1998b)
Inferior colliculus	cat	anesthetized	xtra-cell	headphones	1.5–110 ms (28)	Litovsky & Yin (1998a, 1998b)
Inferior colliculus	cat	anesthetized	xtra-cell	virtual azimuth	2–80 ms (19)	Litovsky *et al.* (1998)
Inferior colliculus	rabbit	awake	xtra-cell	headphones	2–60 ms (6.3)	Fitzpatrick *et al.* (1995)
Inferior colliculus	owl	awake	xtra-cell	free field	1–5 ms	Keller & Takahashi (1996)
Auditory cortex	cat	anesthetized	xtra-cell	virtual 2D free field	48–175 ms (103)	Reale *et al.* (1995; Abstract)
Auditory cortex	rabbit	awake	xtra-cell	headphones	1–100 ms (20)	Fitzpatrick *et al.* (1998; Abstract)

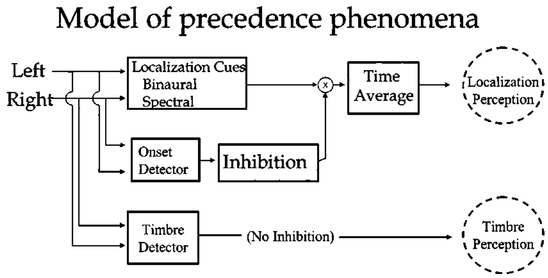

FIG. 13. This model represents a modified version of a model proposed by Zurek (1987) to account for precedence under binaural conditions, in azimuth. In its original form the model contained the localization inputs and an onset detector, which accounted for inhibition of localization information contained in echoes. In its extended form the model also contains a ''timbre detector,'' which lacks inhibition and allows listeners to attend to spectral information contained in echoes. (Upper path reproduced from Zurek, 1987, with permission.)

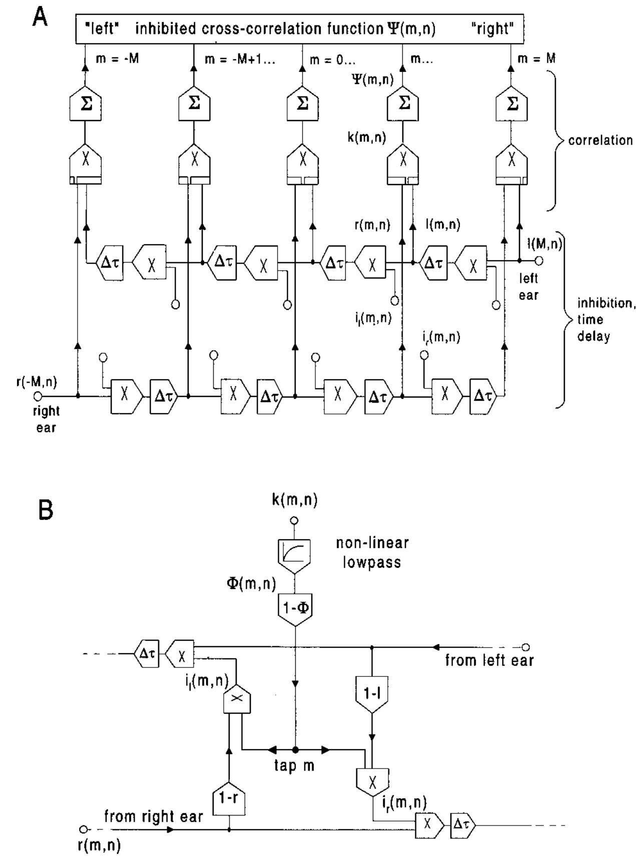

FIG. 14. Inhibited interaural cross-correlation model (Lindemann, 1986b). A: Input signals are represented $[r(-M,n)$ and $l(M,n)]$ which are functions of discrete time variables, *n*. The delayed and attenuated input signals are represented at location *m* as $r(m,n)$ and $l(m,n)$. Input signals are attenuated as they pass along their delay lines, indicated by the symbol *x*. Attenuation is derived such that it increased with level of the contralateral signal. The output at a given location is roughly the correlation between the left and right signals at a given location in the delay line. Hence, wave forms from each side can cancel each other as they pass along the delay line, and it is possible for strong output to inhibit outputs at other locations. (After Lindemann, 1986b; reproduced from Colburn, 1996, with permission.)

factorily. In addition, none of the models can account for phenomena such as the buildup or breakdown of the precedence effect, which are thought to be more cognitive. The discussion here focuses on four models of psychophysical performance (Zurek, 1980, 1987; Lindemann, 1986b; Shinn-Cunningham *et al.*, 1993; Tollin and Henning, 1999), presented in order of their appearance in the literature. As we shall see, these models illustrate fundamentally different approaches and have distinct advantages and disadvantages. In addition, they take a ''bottom-up'' approach with a focus on relatively peripheral effects. At the moment there is no satisfactory model that encompasses basic aspects of the precedence effect. In addition to these psychophysically oriented models, some models of physiological responses (e.g., Wickesberg and Oertel, 1990; Cai *et al.*, 1998a, 1998b) describe neural responses and mechanisms that may be related to some aspects of the precedence effect and they are briefly noted. Finally, we discuss briefly the general question of how much of the precedence phenomena can be understood from a relative peripheral mechanism point of view and how much a more cognitive model is required.

The first model, proposed by Zurek (1980, 1987), is a phenomenological model that is not designed for quantitative predictions but provides an intuitive representation of the basic observations. This model is illustrated in Fig. 13. The upper path of the model was originally proposed by Zurek (1987), who suggested that stimulus onsets (or onsetlike transitions) initiate an inhibition or suppression process that blocks the generation of location information from the ongoing stimulus for a brief period. This upper path thus represents a key aspect of the precedence effect: that the early part of a stimulus leads to suppressive effects on the later part of the stimulus. However, given additional data collected since 1987, we have modified the model by adding the lower path, which emphasized the fact that only localization/lateralization information is suppressed and not other subjective attributes of the stimulus such as loudness, pitch, or timbre. Although this model provides a useful representation of the basic observations, it does not allow quantitative predictions, nor an internal mechanism for the generation of these effects from the stimulus wave forms. A computational model that incorporated these ideas was analyzed by Martin (1997); the model reproduced the basic trends of the data. The general idea of a transient, central inhibitory process that suppressed location information is consistent with the suggestion of Harris *et al.* (1963) as quoted above (end of Sec. I).

The second model, proposed by Lindemann (1986a, 1986b), is an extension of the classic mechanism for sound localization suggested by Jeffress (1948). The Jeffress mechanism, described primarily as a network of coincidence detectors, can be represented as a two-dimensional array of neurons with each neuron's characteristic frequency along one axis and its internal delay along the other axis. Each neuron responds maximally when its immediate inputs are coincident; thus a given neuron responds most vigorously when the stimulus delay is compensated by the neuron's internal delay. For correlated inputs, therefore, interaural delay is translated to a place of maximal stimulation for each frequency. Alternatively, the neural response versus internal delay can be thought of as a cross-correlation function of the bandpass-filtered input wave forms for each characteristic frequency. [For a general discussion of models of binaural processing and the role of the Jeffress-type network in these models, see the review papers of Stern and Trahiotis (1995) and Colburn (1996).] Lindemann (1986a) extended the correlation mechanism to include lateral inhibition along the delay line (see Fig. 14). This inhibition, which includes a static and a dynamic component, suppresses the propagating wave form across the delay lines so that outputs away from the current site of coincidence output are reduced. For stationary stimuli (tones and noise), the model was shown to be consistent with time-intensity trading and with perceptions of image width for decorrelated noises. For nonstationary signals, such as bandpass-filtered transients (Lindemann 1986b), the inhibitory factors lead to a reinforcement of the current locations of activity and a consequent emphasis on the lead sound when there is a lead–lag click pair. In the same paper, this model was also shown to be consistent with several sets of measurements of echo threshold, including the general time scales and the influence of the leading stimulus on the location of the echo. Specifically, the stimulus overlap resulting from the narrow-band filtering of the input leads to the echo location being pulled toward the lead position, even after the echo is perceived as a distinct object. Although this is a powerful model that makes important contributions to our theoretical understanding, the model has not been applied extensively to precedence effect situations, presumably because of its complexity and lack of available implementations. This model is consistent with suggestions of McFadden (1973) in that initial sounds from one direction suppress later sounds from different directions more than sounds from the initial direction. Extensions of this model have been developed by Gaik (1993) to improve the model's response to naturally occurring interaural level differences and by Bodden (1993) to address aspects of the cocktail-party effect.

The third model, proposed by Shinn-Cunningham *et al.* (1993), is a descriptive model that allows the quantitative prediction of lead–lag discrimination data from localization dominance data. The model has been applied to experiments in which the stimulus contains a lead stimulus with ITD τ_1 and a lag stimulus with an ITD τ_2, separated by the lead–lag delay. According to this model, an internal variable is generated which corresponds to the ITD α of the pointer in a localization matching experiment when the lead ITD is τ_1 and the lag ITD is τ_2. The model postulates that

$$\alpha = c*\tau_1 + (1-c)*\tau_2,$$

where c is a non-negative weight that represents the emphasis on the lead relative to the lag stimulus. For example, if the value of $c=0.5$, then lead and lag are equally weighted, there is no precedence effect, and the matching ITD is equal to the average of the lead and lag ITDs. If the value of c is unity ($c=1.0$), then the lead completely dominates the generation of the location and the matching ITD is equal to the lead ITD. (If the value of c were zero, this equation would describe lateral position that is determined by the lag ITD, an extreme form of ''antiprecedence.'') The linear relation assumed here appears to be adequate for almost all position matching data. The dependence of the parameter c on stimulus attributes, such as lead–lag delay and stimulus frequency, allows a simple, quantitative description of localization-dominance results. In addition, by assuming that discrimination of ITD is made on the basis of the combined variable α, and that the separate ITDs are not available, one can predict results of lead–lag discrimination experiments. When these predictions were evaluated (Shinn-Cunningham *et al.*, 1993), the model was shown to explain the quantitative relationship between the results of these two different types of experiments. Specifically, Shinn-Cunningham *et al.* (1993) measured these two tasks in the same listeners. Their model parameters, which were estimated from data obtained with the ''pointer adjustment'' method (see Sec. I B 1), were able to predict results of the ITD discrimination task quite accurately. The disadvantage of this model is that no mechanism is described, so that the model does not *a priori* predict the size of c for different experiments, nor is there a natural extension to other types of experiments or stimuli. Also, as described so far, the model is limited to ITD experiments and to two stimulus components with well-defined lead and lag delays with a single fused image, but the model could be naturally extended to multiple bursts, to angles instead of delays, and to multiple images. This model has been extended to free-field experiments by Litovsky and Macmillan (1994). More recently, Stellmack *et al.* (1998) confirmed the usefulness of this model for interpreting discrimination experiments and obtained similar values for the weighting parameter c.

The fourth model (Tollin, 1998) provides a mechanism for the combination of interaural phase and level information with the time relations of the stimulus components explicitly included in the model. This model has been applied to pairs of click stimuli with a focus on the discrimination of ITD in lead or lag stimulus. In general, estimates of ITD and ILD are combined to generate a decision variable corresponding to lateral position. A notable feature of this model is that, for the cases to which it has been applied, namely the discrimination of wideband clicks, there is no inhibition contained within the model, and yet the model is able to predict that discrimination of the lead click is significantly better than discrimination of the lag click. The basic elements of this model include: a bandpass, rectified cross-correlation function calculation to generate an ITD-based estimate of location; a short-term energy ratio calculation for an ILD-based estimate of location; and a linear weighting of these estimates to generate a final lateral position. A central feature of the model is that the bandpass filter that is used for the estimate of position of a wideband transient stimulus is the filter that is close to 750 Hz (the dominance region) and that is near a peak of the stimulus magnitude spectrum for one ear. This model is also able to predict several ''anomalous localization'' results seen in reports from Wallach *et al.* (1949), Gaskell (1983), and Tollin and Henning (1998). In these cases, the localized direction is opposite to the direction expected from the overall ITD and ILD of the stimulus. The focus on the 750-Hz region and the combination of time and intensity differences for the generation of the lateral position variable and the discrimination decision variable are consistent with the ideas in Gaskell (1983). The model has yet to be applied to a wide class of stimuli and to a wide range of parameters.

To the extent that precedence effects are consistent with the physiological observations discussed in the previous section, models of the physiology should lead to models of precedence. Thus for example, the mechanism suggested by Wickesberg and Oertel (1990) may be involved in monaural processing that influences precedence. Similarly, the models of Cai *et al.* (1998a, 1998b), which show that simple mechanisms of inhibition can successfully describe the ''best-ITD-inhibited'' and ''worst-ITD-inhibited'' neurons described above, may capture aspects of processing that are critical for precedence effects. These physiological models have not been applied to predict the psychophysical performance explicitly and they are not considered further in this review.

It may be noted that the models described in this section are focused on relatively peripheral levels and, as such, are limited to relatively simple, noncognitive mechanisms. However, some aspects of the precedence effects are much more naturally described in cognitive terms, including the buildup and breakdown effects seen with sequences of stimuli. These aspects of precedence illustrate that the peripheral mechanisms included in these models may tell only part of the story. It is appropriate to pursue both types of models, and to keep in mind that a general understanding of precedence effects may require more elaborate interpretations.

VI. CONCLUSION

We have attempted to provide an updated overview of recent psychoacoustical and physiological work on the effects of precedence. By introducing the framework of fusion, localization dominance, and lag-discrimination suppression along with the buildup and breakdown of suppression, we wanted to provide a framework within which to describe the effects of precedence and to help in the integration of data from both psychophysical and physiological experiments. While the motivation for many of the experiments is partly rooted in our desire to understand how the auditory system processes multiple arrays of directional cues in reverberant spaces, many of the studies do not directly address ''real-world'' issues. In fact, the precedence effect as an auditory phenomenon has little to do with realistic acoustic environments. However, we believe that the work in this area can be instrumental in motivating future work in more realistic scenarios, and that it offers a unique window into the auditory system, by viewing auditory processes through human perception, animal behavior and physiology, as well as development. It is probably only through the combined efforts of these various approaches that a full theory of precedence will evolve and useful models will be developed. Since this topic of research is dynamic and ongoing, our paper represents a mere snapshot that will clearly need updating as the field progresses.

ACKNOWLEDGMENTS

The work of Ruth Litovsky is supported by NIDCD Grant (Nos. DC02696 and DC00100) and by ONR-managed MURI Grant (No. Z883401). The work of H. Steven Colburn is supported by NIDCD Grant (No. DC00100). The work of William A. Yost and Sandra Guzman is supported by an NIDCD Program Project Grant (No. DC00293) and by the Air Force Office of Scientific Research. We are grateful to Drs. Rachel Clifton, Pierre Divenyi, Daniel Tollin, and an anonymous reviewer, for their insightful comments on an earlier version of the manuscript. A special thanks are due to Bob Dizon, for many insightful comments and stimulating discussions during the past few years. We would also like to thank Drs. John Brugge and Duck Kim for contributing physiological data collected in their laboratories.

Adams, J., and Mugnaini, E. (**1984**). ''Dorsal nucleus of the lateral lemniscus: A nucleus of GABAergic projection neurons,'' Brain Res. Bull. **13**, 585–590.

Altman, J. A. (**1975**). ''Neurophysiological mechanisms in auditory localization,'' in Brain information services, Soviet research reports (BRI Publication Office, Univ. of California, Los Angeles), Vol. 1, pp. 27–32.

Aoki, S., and Houtgast, T. (**1992**). ''A precedence effect in the perception of interaural cross correlation,'' Hearing Res. **59**, 25–30.

Ashmead, D. H., Clifton, R. K., and Reese, E. P. (**1986**). ''Development of auditory localization in dogs: Single source and precedence effect sounds,'' Dev. Psychobiol. **19**, 91–103.

von Békésy, G. (**1930**). ''Zur Theorie des Hörens: Über das Richtungshören bei einer Zeitdifferenz oder Lautstärkeungleichheit der beidseitigen Schalleinwirkungen [On the theory of hearing: On directional hearing in connection with a time difference or inequality of loudness of the effective sound between the two sides],'' Phys. Z. **31**, 824–838; **31**, 857–868.

Blauert, J. (**1969**). ''Sound localization in the median plane,'' Acustica **22**, 205–213.

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 Edition* (The MIT Press, Cambridge, MA).

Blauert, J., and Col, J. P. (**1992**). ''Irregularities in the precedence effect,'' in *Auditory Physiology and Perception*, edited by Y. Cazals, L. Demaney, and K. Horner (Pergamon, Oxford), pp. 531–538.

Blauert, J., and Divenyi, P. L. (**1988**). ''Spectral selectivity in binaural contralateral inhibition,'' Acustica **66**, 267–274.

Bodden, M. (**1993**). ''Modeling human sound source localization and the cocktail-party-effect,'' Acta Acust. **1**, 43–55.

de Boer, K. (**1940**). ''Three-dimensional sound reproduction,'' Philips Tech. Rev. **5**, 107–115.

Boettcher, F. A., Salvi, R. J., and Saunders, S. S. (**1990**). ''Recovery from short-term adaptation in single neurons in the cochlear nucleus,'' Hearing Res. **48**, 125–144.

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

Cai, H., Carney, L. H., and Colburn, H. S. (**1998a**). ''A model for binaural response properties of inferior colliculus neurons: I. A model with ITD-sensitive excitatory and inhibitory inputs,'' J. Acoust. Soc. Am. **103**, 475–493.

Cai, H., Carney, L. H., and Colburn, H. S. (**1998b**). ''A model for binaural response properties of inferior colliculus neurons: II. A model with ITD-sensitive excitatory and inhibitory inputs and an adaptation mechanism,'' J. Acoust. Soc. Am. **103**, 494–506.

Carney, L. H., and Yin, T. C. T. (**1989**). ''Responses of low-frequency cells in the inferior colliculus to interaural time differences of clicks: Excitatory and inhibitory components,'' J. Neurophysiol. **62**, 144–161.

Clifton, R. K. (**1985**). ''The precedence effect: Its implications for developmental questions,'' in *Auditory Development in Infancy*, edited by S. E. Trehub and B. Schneider (Plenum, New York), pp. 85–99.

Clifton, R. K. (**1987**). ''Breakdown of echo suppression in the precedence effect,'' J. Acoust. Soc. Am. **82**, 1834–1835.

Clifton, R. K., and Freyman, R. L. (**1989**). ''Effect of click rate and delay on breakdown of the precedence effect,'' Percept. Psychophys. **46**(2), 139–145.

Clifton, R. K., and Freyman, R. L. (**1997**). ''The precedence effect: Beyond echo suppression,'' in *Binaural and Spatial Hearing in Real and Virtual Environments*, edited by R. H. Gilkey and T. R. Anderson (Lawrence Earlbaum, Mahwah, NJ), pp. 233–255.

Clifton, R. K., Freyman, R. L., Litovsky, R. Y., and McCall, D. (**1994**). ''Listeners' expectations about echoes can raise or lower echo threshold,'' J. Acoust. Soc. Am. **95**(3), 1525–1533.

Clifton, R. K., Morrongiello, B. A., Kulig, J. W., and Dowd, J. M. (**1981**). ''Newborns' orientation toward sound: Possible implications for cortical development,'' Child Dev. **52**, 833–838.
Colburn, H. S. (**1996**). ''Computational models of binaural processing,'' in *Springer Handbook of Auditory Research: Auditory Computation*, edited by R. R. Fay and A. N. Popper (Springer, New York), pp. 332–400.
Cranford, J. L. (**1982**). ''Localization of paired sound sources in cats: Effects of variable arrival times,'' J. Acoust. Soc. Am. **72**(4), 1309–1311.
Cranford, J. L., and Oberholtzer, M. (**1976**). ''Role of neocortex in binaural hearing in the cat, II: The precedence effect in sound localization,'' Brain Res. **111**, 225–239.
Cranford, J. L., Ravizza, R., Diamond, I. T., and Whitfield, I. C. (**1971**). ''Unilateral ablation of the auditory cortex in the cat impairs complex sound localization,'' Science **172**, 286–288.
David, E. E. (**1959**). ''Comment on the precedence effect,'' in *Proceedings of the 3rd International Congress On Acoustics*, Stuttgart, Vol. 1, pp. 144–146.
Divenyi, P. L. (**1992**). ''Binaural suppression of nonechoes,'' J. Acoust. Soc. Am. **91**(2), 1078–1084.
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**, 3106(A).
Ebata, M., Sone, T., and Nimura, T. (**1968**). ''On the perception of direction of echo,'' J. Acoust. Soc. Am. **44**, 542–547.
Fisher, H. and Freedman, S. J. (**1968**). ''The role of the pinna in auditory localization,'' J. Auditor. Res. **8**, 15–26.
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.
Fitzpatrick, D. C., Kuwada, S., and Kim, D. O. (**1998**). ''Responses of neurons to click-pair stimuli: Auditory nerve to auditory cortex,'' Assoc. Res. Otolaryngology (Abstract; p. 53).
Franssen, N. V. (**1960**). ''Some considerations on the mechanism of directional hearing,'' Ph.D. thesis, Technische Hogeschool, Delft, The Netherlands.
Franssen, N. V. (**1963**). *Stereophony* (Phillips Technical Bibliography, Eindhoven).
Freyman, R. L., Clifton, R. K., and Litovsky, R. Y. (**1991**). ''Dynamic processes in the precedence effect,'' J. Acoust. Soc. Am. **90**, 874–884.
Freyman, R. L., McCall, D. M., and Clifton, R. K. (**1998**). ''Intensity discrimination for precedence effect stimuli,'' J. Acoust. Soc. Am. **103**, 2031–2041.
Gaik, W. (**1993**). ''Combined evaluation of interaural time and intensity differences: Psychoacoustic results and computer modeling,'' J. Acoust. Soc. Am. **94**, 98–110.
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.
Gaskell, H. (**1983**). ''The precedence effect,'' Hearing Res. **12**, 277–303.
Gilkey, R. H., and Anderson, T. R. (Eds.) (**1997**). *Binaural and Spatial Hearing in Real and Virtual Environments* (Lawrence Earlbaum, Mahwah, NJ).
Grantham, D. W. (**1996**). ''Left-right asymmetry in the buildup of echo suppression in normal-hearing adults,'' J. Acoust. Soc. Am. **99**(2), 1118–1123.
Guzman, S. J., and Yost, W. A. (**1999**). ''Forward masking in the precedence effect,'' J. Acoust. Soc. Am. (submitted).
Haas, H. (**1949**). ''The influence of a single echo on the audibility of speech,'' J. Audiol. Eng. Soc. **20**, 145–159, English translation (1972).
Haas, H. (**1951**). ''On the influence of a single echo on the intelligibility of speech,'' Acustica **1**, 48–58.
Hafter, E. R. (**1997**). ''Binaural adaptation and the effectiveness of a stimulus beyond its onset,'' in *Binaural and Spatial Hearing in Real and Virtual Environments*, edited by R. H. Gilkey and T. R. Anderson (Lawrence Earlbaum, Mahwah, NJ), pp. 211–232.
Hafter, E. R., and Buell, T. N. (**1990**). ''Restarting the adapted binaural system,'' J. Acoust. Soc. Am. **88**(2), 806–812.
Hafter, E. R., Buell, T. N., and Richards, V. (**1988**). ''Onset-coding in lateralization: Its form, site, and function,'' in *Auditory Function: Neurobiological Bases of Hearing*, edited by G. M. Edelman, W. E. Gall, and W. M. Cowan (Wiley, New York), pp. 647–676.
Hafter, E. R., Dye, R. H., and Wenzel, E. (**1983**). ''Detection of interaural differences of intensity in trains of high-frequency clicks as a function of interclick interval and number,'' J. Acoust. Soc. Am. **73**, 1708–1713.
Hafter, E. R., Saberi, K., Jensen, E. R., and Briolle, F. (**1992**). ''Localisation in an echoic environment,'' in *Auditory Physiology and Perception*, edited by Y. Cazals, L. Demaney, and K. Horner (Pergamon, Oxford), pp. 555–561.
Harris, G. G., Flanagan, J. L., and Watson, B. J. (**1963**). ''Binaural interaction of a click with a click pair,'' J. Acoust. Soc. Am. **35**, 672–678.
Hartmann, W. M., and Rakerd, B. (**1989**). ''Localization of sound in rooms IV: The Franssen effect,'' J. Acoust. Soc. Am. **86**(4), 1366–1373.
Hochster, M. E., and Kelly, J. B. (**1981**). ''The precedence effect and sound localization by children with temporal lobe epilepsy,'' Neuropsychologia **19**, 49–55.
Jeffress, L. A. (**1948**). ''A place theory of sound localization,'' J. Comp. Physiol. Psychol. **41**(1), 35–39.
Kaltenbach, J. A., Meleca, R. J., Falzarano, P. R., Myers, S. F., and Simpson, T. H. (**1993**). ''Forward masking properties of neurons in the dorsal cochlear nucleus: Possible role in the process of echo suppression,'' Hearing Res. **67**, 35–44.
Keller, C. H., and Takahashi, T. T. (**1996**). ''Responses to simulated echoes by neurons in the barn owl's auditory space map,'' J. Comp. Physiol. **178**(4), 499–512.
Kelly, J. B. (**1974**). ''Localization of paired sound sources in the rat: Small time differences,'' J. Acoust. Soc. Am. **55**, 1277–1284.
Kuwada, S., Batra, R., and Stanford, T. R. (**1989**). ''Monaural and binaural response properties of neurons in the inferior colliculus of the rabbit: Effects of sodium pentobarbital,'' J. Neurophysiol. **61**, 269–282.
Leakey, D. M. (**1959**). ''Some measurements of the effects of interchannel intensity and time differences in two channel sound systems,'' J. Acoust. Soc. Am. **31**, 977–986.
Leakey, D. M., and Cherry, E. C. (**1957**). ''Influence of noise upon the equivalence of intensity differences and small time delays in two-loudspeaker systems,'' J. Acoust. Soc. Am. **29**, 284–286.
Liberman, M. C. (**1978**). ''Auditory nerve responses from cats raised in a low-noise chamber,'' J. Acoust. Soc. Am. **63**, 442–455.
Lindemann, W. (**1986a**). ''Extension of a binaural cross-correlation model by contralateral inhibition. I. Simulation of lateralization for stationary signals,'' J. Acoust. Soc. Am. **80**, 1608–1622.
Lindemann, W. (**1986b**). ''Extension of a binaural cross-correlation model by contralateral inhibition. II. The law of the first wave front,'' J. Acoust. Soc. Am. **80**, 1623–1630.
Litovsky, R. Y. (**1997**). ''Developmental changes in the precedence effect: Estimates of Minimal Audible Angle,'' J. Acoust. Soc. Am. **102**, 1739–1745.
Litovsky, R. Y., Cransten, B. R., and Delgutte, B. (**1998**). ''Neural correlates of the precedence effect in the inferior colliculus: Effect of localization cues,'' Assoc. Res. Otolaryngology (Abstract, p. 40).
Litovsky, R. Y., Hawley, M. L., and Colburn, H. S. (**1997a**). ''Measurement of precedence in monaural listeners,'' Meeting of the American Speech and Hearing Association, Boston, MA.
Litovsky, R. Y., Rakerd, B., Yin, T. C. T., and Hartmann, W. M. (**1997b**). ''Psychophysical and physiological evidence for a precedence effect in the median saggital plane,'' J. Neurophysiol. **77**, 2223–2226.
Litovsky, R. Y. (**1998**). ''Physiological studies on the precedence effect in the inferior colliculus of the kitten,'' J. Acoust. Soc. Am. **103**, 3139–3152.
Litovsky, R. Y. and Colburn, H. S. (**1998**). ''Precedence effects in the azimuthal and sagittal planes,'' Assoc. Res. Otolaryngology Abstracts **21**, p. 53.
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. (**1999b**). ''Physiological studies of the precedence effect in the inferior colliculus of the cat: II. Neural mechanisms,'' J. Neurophysiol. **80**, 1302–1316.
Litovsky, R. Y., Dizon, R. M., and Colburn, H. S. (**1999**). ''Studies of the precedence effect in the median-sagittal and azimuthal planes in a virtual acoustic space,'' J. Acoust. Soc. Am. (submitted).
Litovsky, R. Y., and Ashmead, D. H. (**1997**). ''Developmental aspects of binaural and spatial hearing,'' in *Binaural and Spatial Hearing in Real and Virtual Environments*, edited by R. H. Gilkey and T. R. Anderson (Lawrence Earlbaum, Mahwah, NJ), pp. 571–592.
Litovsky, R. Y., and Macmillan, N. A. (**1994**). ''Sound localization precision under conditions of the precedence effect: Effects of azimuth and standard stimuli,'' J. Acoust. Soc. Am. **96**(2) Pt. 1, 752–758.
Lochner, J. P. A., and Burger, J. F. (**1958**). ''The subjective masking of

short time delayed echoes, their primary sounds, and their contribution to the intelligibility of speech,'' Acustica **8**, 1–10.

McFadden, D. (**1973**). ''Precedence effect and auditory cells with long characteristic delays,'' J. Acoust. Soc. Am. **54**, 528–530.

Martin, K. D. (**1997**). ''Echo suppression in a computational model of the precedence effect,'' presented at the 1997 IEEE Mohonk Workshop on Applications of Signal Processing to Acoustics and Audio, New Paltz, NY.

Meyer, E., and Schrodder, G. R. (**1952**). ''Über den Einfluss von Schallrückwürfen auf Richtungslokalisation und Lautstärke bei Sprache [On the influence of reflected sound on directional localization and loudness of speech],'' Nachr. Akad. Wiss. Goett. II, Math.-Phys. Kl. **6**, 31–42.

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

Moore, B. C. J. (**1996**). *An Introduction to the Psychology of Hearing* (Academic, New York).

Moore, J. K., and Moore, R. Y. (**1987**). ''Glutamic-acid decarboxylase-like immunoreactivity in brain-stem auditory nuclei of the rat,'' J. Comp. Neurol. **260**(2), 157–174.

Morrongiello, B. A., Kulig, J. W., and Clifton, R. K. (**1984**). ''Developmental changes in auditory temporal perception,'' Child Dev. **55**, 461–471.

Muncey, R. W., Nickson, A. F. B., and Dubout, P. (**1953**). ''The acceptability of speech and music with a single artificial echo,'' Acustica **3**, 168–173.

Oliver, D. L., Beckius, G. E., and Ostapoff, E. M. (**1994**). ''Connectivity of neurons in identified auditory circuits studied with transport of dextram and microspheres plus intracellular injection of Lucifer Yellow,'' J. Neurosci. Methods **53**, 23–27.

Parham, K., Zhao, H. B., and Kim, D. O. (**1996**). ''Responses of auditory nerve fibers to unanesthetized decerebrate cat to click pairs as simulated echoes,'' J. Neurophysiol. **76**, 17–28.

Parham, K., Zhao, H. B., Ye, Y., and Kim, D. O. (**1998**). ''Responses of anteroventral cochlear nucleus neurons of the unanesthetized decerebrate cat to click pairs as simulated echoes,'' Hearing Res. **125**, 131–146.

Perrott, D. R., and Pacheco, S. (**1989**). ''Minimal auditory angle thresholds for broadband noise as a function of delay between the onset of the lead and lag signals,'' J. Acoust. Soc. Am. **85**, 2669–2672.

Perrott, D. R., Marlborough, K., Merrill, P., and Strybel, T. Z. (**1989**). ''Minimum audible angle thresholds obtained under conditions in which the precedence effect is assumed to operate,'' J. Acoust. Soc. Am. **85**, 282–288.

Perrott, D. R., Strybel, T. Z., and Manligas, C. L. (**1987**). ''Conditions under which the Haas precedence effect may or may not occur,'' J. Audit. Res. **27**, 59–72.

Populin, I. C., and Yin, T. C. T. (**1998**). ''Behavioral studies of sound localization in the cat,'' J. Neurosci. **18**, 2147–2160.

Reale, R. A., Brugge, J. F., and Hind, J. E. (**1995**). ''Encoding of stimulus direction by AI neurons in the cat: Effects of sounds arriving at different times and from different directions,'' Neurosci. Abs. **21**, 667.

Rakerd, B., and Hartmann, W. M. (**1985**). ''Localization of sound in rooms, II: The effects of a single reflecting surface,'' J. Acoust. Soc. Am. **78**(2), 524–533.

Rakerd, B., Hsu, J., and Hartmann, W. M. (**1997**). ''The Haas effect with and without binaural differences,'' J. Acoust. Soc. Am. **101**(5), 3083.

Saberi, K., and Perrott, D. R. (**1990**). ''Lateralization thresholds obtained under conditions in which the precedence effect is assumed to operate,'' J. Acoust. Soc. Am. **87**(4), 1732–1737.

Saberi, K. (**1996**). ''Observer weighting of interaural delays in filtered impulses,'' Percept. Psychophys. **58**(7), 1037–1046.

Saintmarie, R. L., Ostapoff, E. M., Morest, D. K., and Wenthold, R. J. (**1989**). ''Glycene-immunoreactive projection of the cat lateral superior olive: Possible role in midbrain ear dominance,'' J. Comp. Neurol. **279**, 382–396.

Scharf, B. (**1974**). ''Localization of unlike tones from two loudspeakers,'' in *Sensation and measurement: papers in honor of S.S. Stevens*, edited by H. R. Moskowitz, B. Scharf, and J. C. Stevens (Reidel, Dordrecht), pp. 309–314.

Schubert, E. D., and Wernick, J. (**1969**). ''Envelope versus microstructure in the fusion of dichotic signals,'' J. Acoust. Soc. Am. **45**, 1525–1531.

Shaw, A. G. (**1974**). ''The external ear,'' in *Handbook of Sensory Physiology*, edited by W. D. Keidel and W. D. Neff (Springer-Verlag, New York), Vol. V/1, pp. 455–490.

Shinn-Cunningham, B. G., Zurek, P. M., Durlach, N. I., and Clifton, R. K. (**1995**). ''Cross-frequency interactions in the precedence effect,'' J. Acoust. Soc. Am. **98**(1), 164–171.

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**(5), 2923–2932.

Shore, S. E. (**1995**). ''Recovery of forward-masked responses in ventral cochlear nucleus neurons,'' Hearing Res. **82**, 31–43.

Snow, W. (**1954a**). ''Basic principles of stereophonic sound,'' J. Soc. Motion Picture Television Eng. **61**, 567–589.

Snow, W. B. (**1954b**). ''Effect of arrival time on stereophonic localization,'' J. Acoust. Soc. Am. **26**, 1071–1074.

Stellmack, M. A., Dye, R. H., and Guzman, S. J. (**1999**). ''Observer weighting of binaural information in source and echo clicks,'' J. Acoust. Soc. Am. **105**, 377–387.

Stern, R. M., and Trahiotis, C. (**1995**). ''Models of binaural interaction'' in *Handbook of Perception and Cognition*, edited by B. C. J. Moore (Academic, New York), Vol. 6, pp. 347–386.

Thurlow, W. R., and Parks, T. E. (**1961**). ''Precedence-suppression effects for two click sources,'' Percept. Mot. Skills **13**, 7–12.

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*, I-Lucca, Italy, 1–12 July 1998, edited by S. Greenberg and M. Slaney, pp. 77–82.

Tollin, D. J., and Henning, G. B. (**1998**). ''Some aspects of the lateralization of echoed sound in man. I: Classical interaural delay-based precedence.'' J. Acoust. Soc. Am. **104**, 3030–3038.

Tollin, D. J., and Henning, G. B. (**1999**). ''Some aspects of the lateralization of echoed sound in man. II: The role of stimulus spectrum,'' J. Acoust. Soc. Am. **105**, 838–849.

Wallach, H., Newman, E. B., and Rosenzweig, M. R. (**1949**). ''The precedence effect in sound localization,'' Am. J. Psychol. **LXII**(3), 315–336.

Warncke, H. (**1941**). ''The fundamentals of room-related stereophonic reproduction in sound films,'' Akust. Zh. **6**, 174–188.

Whitfield, I. C. (**1978**). ''Auditory cortical lesions and the precedence effect in a four-choice situation,'' J. Physiol. (London) **289**, 81.

Whitfield, I. C., Cranford, J., Ravizza, R., and Diamond, I. T. (**1972**). ''Effects of unilateral ablation of auditory cortex in cat on complex sound localization,'' J. Neurophysiol. **35**, 718–731.

Wickesberg, R. E. (**1996**). ''Rapid inhibition in the cochlear nuclear complex of the chinchilla,'' J. Acoust. Soc. Am. **100**(3), 1691–1702.

Wickesberg, R. E., and Oertel, D. (**1990**). ''Delayed, frequency-specific inhibition in the cochlear nuclei of mice: A mechanism for monaural echo suppression,'' J. Neurosci. **10**(6), 1762–1768.

Wyttenbach, R. A., and Hoy, R. R. (**1993**). ''Demonstration of the precedence effect in an insect,'' J. Acoust. Soc. Am. **94**, 777–784.

Yang, X., and Grantham, D. W. (**1997a**). ''Echo suppression and discrimination suppression aspects of the precedence effect,'' Percept. Psychophys. **59**, 1108–1117.

Yang, X., and Grantham, D. W. (**1997b**). ''Cross-spectral and temporal factors in the precedence effect: Discrimination suppression of the lag sound in free-field,'' J. Acoust. Soc. Am. **102**, 2973–2983.

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.

Yost, W. A., Patterson, R., and Sheft, S. (**1996**). ''A time domain description of iterated rippled noise,'' J. Acoust. Soc. Am. **99**(2), 1066–1078.

Yost, W. A., and Guzman, S. J. (**1996**). ''Auditory processing of sound sources: Is there an echo in here?'' Curr. Directions Psychol. Sci. **5**(4), 125–131.

Yost, W. A., and Hafter, E. R. (**1987**). ''Lateralization,'' in *Directional Hearing*, edited by W. A. Yost and G. Gourevitch (Springer-Verlag, New York), pp. 49–84.

Yost, W. A., and Soderquist, D. R. (**1984**). ''The precedence effect: Revisited,'' J. Acoust. Soc. Am. **76**(5), 1377–1383.

Yost, W. A., Mapes-Riordan, D., and Guzman, S. J. (**1997**). ''The relationship between localization and the Franssen effect,'' J. Acoust. Soc. Am. **101**(5), 2994–2997.

Zurek, P. M. (**1980**). ''The precedence effect and its possible role in the avoidance of interaural ambiguities,'' J. Acoust. Soc. Am. **67**(3), 952–964.

Zurek, P. M. (**1987**). ''The precedence effect,'' in *Directional Hearing*, edited by W. A. Yost and G. Gourevitch (Springer-Verlag, New York), pp. 85–105.