# Detection of spatial cues in linear and logarithmic frequency-modulated sweeps

I-HUI HSIEH

National Central University, Jhongli City, Taiwan

AND

KOUROSH SABERI University of California, Irvine, California

Frequency- and amplitude-modulated (FM and AM) sounds are the building blocks of complex sounds. In the present study, we investigated the ability of human observers to process spatial information in an important class of FM sounds: broadband directional sweeps common in natural communication signals such as speech. The stimuli consisted of linear or logarithmic unidirectional FM pulses that swept either up or down in frequency at various rates. Spatial localization thresholds monotonically improved as sweep duration decreased and as sweep rate increased, but no difference in performance was observed between logarithmic and linear or between up-and down-frequency sweeps. Counterintuitive reversals in localization were observed which suggested that the localization of high-frequency sweeps may be strongly dominated by amplitude information even in situations in which one might consider timing cues to be critical. Implications of these findings for the localization of complex sounds are discussed.

Many naturally occurring sounds are modulated in amplitude or frequency; important examples include speech and other conspecific communication signals in mammals, birds, marine species, and even insects (Bailey, Greenfield, & Shelly, 1993; Brillet & Paillette, 1991; Coscia, Phillips, & Fentress, 1991; Dankiewicz, Helweg, Moore, & Zafran, 2002; Dear, Simmons, & Fritz, 1993; Fant, 1970; Huber & Thorson, 1985; Klump & Langemann, 1992; Pickett, 1980; Robisson, Aubin, & Bremond, 1993; Ryan & Wilczynskin, 1988; Saberi & Perrott, 1999; Sabourin, Gottlieb, & Pollack, 2008; Simmons, 1979). Because amplitude- and frequency-modulated (AM and FM) sounds are the building blocks of complex sounds, understanding how the auditory system encodes these signals has important practical and theoretical implications (Kay, 1982; Moore & Sek, 1992; Saberi, 1998). The peripheral mechanisms of AM encoding are relatively well understood, requiring a tracking of changes in firing rates of primary auditory afferents in response to the modulation envelope (Javel, 1980; Langner & Schreiner, 1988; Smith & Brachman, 1980). The mechanisms for coding FM signals are not as clear, because the FM envelope is flat. One biological solution is to monitor amplitude fluctuations at the outputs of cochlear filters as the instantaneous frequency of the FM sweeps through the passband of these filters (Blauert, 1981; Saberi, 1998; Saberi & Hafter, 1995). This FM-to-AM transduction is extremely useful, because not only does it provide a neurally economical solution to encoding both AM and FM

sounds, but, from a theoretical standpoint, it also facilitates analysis of such important auditory theories as the duplex theory of binaural localization.

The duplex theory, formulated in a seminal paper by Lord Rayleigh at the turn of the 20th century, proposed that humans use interaural time differences (ITDs) at low frequencies (below 1.5 kHz) and interaural level differences (ILDs) at high frequencies to localize sounds along the azimuthal plane (Rayleigh, 1907). In the early to mid-1970s, the viability of this theory was challenged by studies that demonstrated use of ITDs in envelopes of highfrequency AM sinusoidal carriers (Henning, 1974, 1980; McFadden & Moffitt, 1977; McFadden & Pasanen, 1975, 1976; Nuetzel & Hafter, 1976). Neurophysiological findings, in addition, showed that some binaurally sensitive neurons can encode ITDs in high-frequency AM sounds by phase locking to their slowly modulating amplitude envelopes (Crow, Langford, & Moushegian, 1980). Later psychophysical studies showed that subjects can also use ITDs in narrowband sinusoidal FM (SFM) signals at high frequencies (Blauert, 1981; Henning, 1980; Nuetzel & Hafter, 1981; Saberi, 1998; Saberi & Hafter, 1995).

The present study had two specific goals: The first was to empirically determine the binaural system's ability to process an important class of FM signals not previously examined in spatial hearing research. This class of stimuli—broadband directional FM sweeps—is a common feature of many types of natural sounds. Although directional sweeps have been used extensively in animal

I.-H. Hsieh, ihsieh@ncu.edu.tw

neurophysiological studies (Brimijoin & O'Neill, 2005; Fuzessery, Richardson, & Coburn, 2006; Mendelson & Grasse, 1992; Mendelson, Schreiner, Sutter, & Grasse, 1993; Nelken & Versnel, 2000; Razak & Fuzessery, 2006, 2007), human developmental research (Colombo & Horowitz, 1986), studies of tonal languages (Luo, Boemio, Gordon, & Poeppel, 2007), and studies of music perception (d'Alessandro, Rosset, & Rossi, 1998), in no prior study has the spatial sensitivity of the auditory system to these sounds been investigated.<sup>1</sup> The second goal of this study was to investigate theoretical issues related to the localization of high-frequency complex sounds. Several

issues are addressed, including how extensions to the duplex theory on the use of envelope ITD cues at high frequencies are predicted from an FM-to-AM transduction and the extent to which envelope ITD processing may be accounted for by the presence of ILD cues extracted from cross-channel spectral disparity.

Specifically, we examined ITD sensitivity to directional linear and logarithmic FM sweeps as a function of sweep direction, type, and rate. We used both an up and a down sweep because previous work has suggested the existence of neurons selective to sweep direction (Mendelson & Cynader, 1985; Mendelson & Grasse, 1992; Nelken & Versnel, 2000). We investigated binaural sensitivity to both linear and logarithmic sweeps because envelope fluctuations at the outputs of auditory filters will have different rise/decay slopes, depending on sweep type, and because most previous work in this area has included only linear sweeps in spite of the logarithmically organized peripheral and cortical tonotopy (Clopton, Winfield, & Flammino, 1974; M. C. Liberman, 1982; Müller, von Hünerbein, Holdis, & Smolders, 2005; Romani, Williamson, & Kaufman, 1982). Finally, we examined ITD sensitivity as a function of sweep duration, ranging from 0.5 msec to 1 sec and, surprisingly, found that durations exceeding 10 msec produced near-chance performance. A number of other unexpected observations also emerged, which contrast with predictions of a cross-correlation model of binaural interaction that utilizes fine structure and envelope interaural delay cues from FM-to-AM transformation at the outputs of simulated cochlear filters.

## **EXPERIMENT 1** Detection of Interaural Delays in FM Sweeps

## Method

**Subjects**. Three normal-hearing adults participated in this experiment. Two were male and 1 female. All of them were highly experienced in psychoacoustic experiments, and were each practiced the various conditions of the experiment for 2 h prior to data collection.

**Stimuli**. The stimuli were generated using MATLAB software (MathWorks) on a Dell PC (Dimension 8400) and were presented at a rate of 44.1 kHz through 16-bit digital-to-analog converters (Creative Sound Blaster Audigy 2ZS) and through Sennheiser headphones (HD 470) in a double-walled steel acoustically isolated chamber (Industrial Acoustics Company). The linear FM sweeps were generated from Equation 1:

$$X(t) = \sin\left[2\pi f_s(t + \text{ITD}) + \pi \frac{f_e - f_s}{T_s}(t + \text{ITD})^2\right], \quad (1)$$

where  $f_s$  and  $f_e$  represent the starting and ending sweep frequencies in hertz, and  $T_s$  is stimulus duration. ITD was set to zero in one channel and to the desired interaural delay in the other. Note, however, that a positive ITD in this equation results in an increase in start frequency and hence represents a waveform leading in time. The sweep's instantaneous frequency is

$$f_i(t) = f_s + (t + \text{ITD}) \frac{\left(f_e - f_s\right)}{t_s}.$$
(2)

The logarithmic FM sweeps were generated from Equation 3:

$$Y(t) = \sin\left(2\pi \frac{f_s}{\frac{\ln(f_e/f_s)}{T_s}} \left\{ \exp\left[\frac{\ln(f_e/f_s)}{T_s}(t + \text{ITD})\right] - 1 \right\} \right\}, \quad (3)$$

with an instantaneous frequency,

$$g_{i}(t) = e^{\ln(f_{s}) + (t + \operatorname{ITD})\left[\ln(f_{e}) - \ln(f_{s})\right]/T_{s}}.$$
(4)

In the present study, we selected a starting frequency of 3 kHz, randomized by 10% on each presentation (3-3.3 kHz), and an end frequency of 8 kHz (+10% rove; i.e., 5-kHz sweep). We selected a high starting frequency to avoid carrier interaural phase effects, which are dominant at frequencies below 1.5 kHz (Mills, 1960, 1972; Rayleigh, 1907; Yost & Hafter, 1987). This allowed a test of the hypothesized encoding of envelope interaural delays across frequency-matched filters as the FM sweeps through their passband. The frequency range of 3-8 kHz is also a critical range for a variety of other spatial processes-for example, the externalization of HRTF-filtered sounds (Butler, 1975; Butler & Belendiuk, 1977; Wightman & Kistler, 1989a, 1989b). All waveforms had simultaneous onsets and offsets in the two channels to prevent the use of interaural rise/decay envelope cues. Delays between the left and right channels were checked for accuracy with a dual-channel digital storage oscilloscope (Tektronix, Model TDS210).

Six stimulus durations, and thus six sweep rates, were used. These durations were 0.5, 1, 5, 10, 25, and 50 msec, producing extremely rapid sweep rates ranging from 0.1 to 10 kHz/msec for linear sweeps and from 0.0283 to 2.83 octaves/msec for log sweeps  $[log_2(8/3)/duration]$ .<sup>2</sup> We selected these values on the basis of initial pilot work, which surprisingly showed that FM sweeps with durations exceeding 50 msec could not be lateralized on the basis of an interaural delay.

To reduce the effects of interaural level resulting from the headphone transfer function when the instantaneous frequency in one channel lags that in the other (or resulting from a difference between the transfer functions of the two channels of the headphones), we filtered each stimulus by scaling it with the inverse of its Hilbert envelope, measured at the output of each channel separately at a sampling rate of 44.1 kHz, using a 6-cc coupler, a 0.5-in. microphone (Brüel & Kjær Model 4189), a conditioning amplifier (Nexus, Brüel, & Kjær), and a 16-bit analog-to-digital converter (Creative Sound Blaster Audigy 2ZS). Analysis of the recorded waveform showed that the amplitude envelopes were flattened to within 0.5 dB of the mean level across the entire range of frequency sweeps tested. In addition to using the Hilbert envelopes to flatten the waveforms at the output of the headphones, we also tested flattening the headphone transfer functions by using a pair of complementary 512-point Golay codes (Zhou, Green, & Middlebrooks, 1992), with which we inverse filtered the FM sweeps. Although both methods provided relatively flat amplitude envelopes, the Hilbert transform method produced slightly better results. The level of a 1-sec sweep was set to 75 dB SPL (A weighted). No other changes were made to stimulus levels. Except for the 0.5- and 1-msec stimulus durations, all stimuli had linear rise/decay times of 1 msec.

**Procedure**. The experiment was run in a block design in which the sweep's duration, direction (up or down in frequency), and type (linear or log) were held constant within a run. Each subject completed six runs for each of 24 conditions in a random-block design in which 1 of 24 conditions was randomly selected without replacement on each run until a full set of 24 conditions was completed. This process was then continued five more times. Each run consisted of 50 trials in a two-interval forced-choice (2IFC), two-down oneup adaptive design, in which the subjects' 70.7% correct-response threshold was tracked (Levitt, 1971; Wetherill & Levitt, 1965).

On the first interval of each trial, the dichotic sweep led to one randomly selected ear by a specific ITD, and in the second interval, it led to the other ear by the same magnitude of ITD. The subjects' task was to identify the order of presentation of the sweeps (i.e., left leading, then right, or right leading, then left). Perceptually, this is equivalent to determining whether the two intracranial auditory images in the two intervals of the trial were heard on the left, then the right, or the right, and then on the left. The subjects then pressed either a left or a right key to respond (a left key response meant that they perceived the sound order as right to left). Visual feedback was provided after each trial. Pilot work showed that certain conditions produced percepts of lateralization that were clearly opposite to that predicted from the interaural delay. We will discuss these counterintuitive reversals extensively in the description of Experiments 2A and 2B and in the General Discussion section. The subjects were instructed to use the feedback in any manner they wished to obtain the lowest possible threshold. The feedback was provided by two methods: First, a plot of the staircase response (ITD as a function of trial number) was shown on the monitor with a trial-by-trial update. Second, the subjects were told, via text on the monitor, which interval contained the leading interaural delay. The subjects typically reported that they relied more heavily on the staircase display as feedback and that adjusting their response strategy on the basis of response feedback was quite easy.

The initial value of the total interaural delay on each run was  $1,500 \ \mu$ sec (i.e.,  $750 \ \mu$ sec in each interval). Two successive correct responses led to a reduction of the total interaural delay by a step size of 0.15 log units (Saberi, 1995b). An incorrect response led to an increase in ITD by the same step size. The threshold on each run was estimated as the average of the stimulus values at track reversal points. The first three or four reversals from each run were discarded, and the threshold was estimated as the average of the remaining even number of reversals. Usually, four to eight reversals went into the calculation of each threshold.

### **Results and Discussion**

Figures 1A and 1B show the results from the linear sweep conditions, and Figures 1C and 1D show results from the logarithmic sweep conditions. Figures 1A and 1C show results from the up sweeps, and Figures 1B and 1D show results from the down sweeps. The lower and upper axes represent sweep duration and rate, respectively,<sup>3</sup> and the horizontal line near 1,200  $\mu$ sec represents chance performance derived from a 5,000-run Monte Carlo simulation. The data are averaged across the 3 observers. Error bars represent 1 *SD*. There was a clear effect of sweep rate/duration, with substantial improvement at higher rates. However, there is no significant difference in thresholds as a function of sweep conditions overlaid, without error bars, to facilitate visual comparison.

A three-way (6 × 2 × 2) repeated measures ANOVA on the data of Figures 1A–1D showed a significant effect of sweep rate/duration [F(5,10) = 7.34, p < .005], but no significant effect of sweep direction [F(1,2) =10.07, n.s.] or sweep type [linear vs. log; F(1,2) = 0.87, n.s.]. There was also no duration × direction interaction [F(5,10) = 1.22, n.s.], no duration × sweep type interaction [F(5,10) = 1.68, n.s.], no sweep direction  $\times$  type interaction [F(1,2) = 1.59, n.s.], and no significant three-way interaction between duration, type, and direction [F(5,10) = 2.55, n.s.].

A surprising result is the near-chance threshold observed for longer stimulus durations (25 and 50 msec), which are relatively fast sweeps across the 5-kHz range (100 kHz/sec for the 50-msec linear sweep). In pilot tests, we found similarly high thresholds for longer sweep durations up to 1 sec. Only when duration was reduced to 5 msec or shorter did we observe thresholds in the range of those reported for SFM stimuli. Nearly all previous studies of spatial sensitivity to FM sounds have used narrowband sinusoidal modulators instead of directional sweeps across a broad region of the spectrum. It may, nonetheless, be instructive to compare the thresholds obtained in the present study with those from SFM signals to gain better insight into our findings. Prior research has shown that the best ITD sensitivity for an SFM signal is observed for rates of 200-400 Hz and carriers of 3 or 4 kHz (Saberi, 1998). Figure 2 shows the instantaneous frequency deviation of a 300-Hz sinusoidal modulator (shown by the solid line) and its time derivative (broken line). The sinusoidal modulator has an instantaneous frequency of  $f(t) = f_m \sin(2\pi f_m t)$  and a rate of change in frequency df/dt $dt = 2\pi f_m^2 \cos(2\pi f_m t)$ , where  $f_m$  is the modulation frequency (for a fully modulated waveform, this is equal to peak frequency deviation). Note that the rate of change in frequency, which peaks when instantaneous frequency deviation crosses zero, is approximately 500 kHz/sec, or a factor of five higher than the constant rate associated with the 50-msec linear sweep and equal to that associated with the 10-msec linear sweep. On average, the 10-msec sweep produced thresholds of ~800  $\mu$ sec, which is substantially higher than that reported for SFM stimuli under an equivalent maximum rate of change in frequency (200- $\mu$ sec thresholds for a 3-kHz carrier sinusoidally modulated at 300 Hz). This difference (i.e., lower thresholds for SFM stimuli than for FM sweeps) may in part be because an SFM waveform sweeps through a restricted range of frequencies hundreds of times per second and in part because our unidirectional sweep stimuli pass through a range of frequencies above 5 kHz. SFM thresholds substantially increase with increasing carrier frequency, possibly because of wider filter bandwidths at higher frequencies (see Figure 3 of Saberi, 1998). Note, however, that the shortest duration stimuli, which also sweep from 3 to 8 kHz, produce thresholds in the order of 200  $\mu$ sec, similar to the best thresholds obtained for SFM stimuli.

## EXPERIMENTS 2A AND 2B Control Conditions

Two additional control experiments were conducted to determine the causes of counterintuitive reversals in lateralization, noted in the Method section of Experiment 1. These reversals, which are fully analyzed in the General Discussion section, suggest that interaural level cues in the 3- to 4-kHz region of a dichotic FM sweep dominate the percept of lateralization, even in the presence of other



Figure 1. Interaural delay thresholds averaged across the 3 observers as a function of FM sweep duration (bottom axis) and rate of change in frequency (top axis). Panels A and B show the data from linear frequency-sweep conditions, and C and D show those from logarithmic sweeps. Panels A and C show the data from up sweeps (3–8 kHz) and panels B and D from down sweeps (8–3 kHz). The horizontal line near 1200  $\mu$ sec shows chance performance obtained from a 5,000-run Monte Carlo simulation. Panel E plots all four functions together, with the error bars excluded to facilitate visual comparison.



Figure 2. Instantaneous change in frequency (black, left axis) from a sinusoidal FM modulated at 300 Hz, and its instantaneous rate of change in frequency (right axis; gray in print, red online).

conflicting cues. Prior research has shown no significant difference between ILD thresholds for pure tones in the 3-kHz regions and those for tones at higher frequencies, near 8-10 kHz (Grantham, 1984; Mills, 1960; Yost, 1981; Yost & Dye, 1988). These findings, however, are based on isolated tones in the absence of competing information. The purpose of the first control experiment (Experiment 2A) was to determine whether ILD cues in the 2- to 4-kHz region dominate conflicting ILD cues at higher frequencies when they are simultaneously present in a complex broadband sound. Such a confirmation has implications for the lateralization of interaurally delayed FM sweeps, as is analyzed in the Discussion section of Experiment 2A. A second control experiment (Experiment 2B) was also conducted to ensure that low-frequency ITD cues below 1 kHz did not contribute to the lateralization of brief FM pulses.

## Experiment 2A: Spectral Dominance in ILD Detection

#### Method

**Subjects**. Three normal-hearing adults participated in this experiment. Two of the subjects had participated in Experiment 1, and 1 was a new subject. This latter subject was highly experienced in psychoacoustic experiments and practiced the various conditions of the experiment for 1 h prior to data collection.

**Procedures.** The procedure and apparatus were the same as those in Experiment 1, with the following changes. In Part 1, we presented the same type of stimuli used in Experiment 1 in a 2IFC mixedblock design without feedback to objectively verify the conditions under which reversals of lateralization occurred. The stimulus was a 1-msec linear sweep with a fixed ITD of 650  $\mu$ sec. Two types of FM sweeps (up or down) were randomly presented within the same block, with the interaural delay favoring the right or left ear. As in Experiment 1, the subjects' task was to determine whether the order of perceived locations of stimuli in the two intervals was left then right or vice versa. Each of the 3 subjects completed four runs of 50 trials each. The purpose was to determine whether the perceived lateral positions were consistent with the lead interaural delay or opposite to this delay (i.e., reversal).

In Part 2, we investigated whether an ILD cue in the low-frequency regions of a noise band dominates an equal but opposite ILD cue in a higher frequency region within the same noise band. We generated a 500-msec burst of Gaussian noise, filtered between 2 and 9.5 kHz. The noise level in the 2- to 4.5-kHz region was higher in one ear than in the other ear by 6 dB, whereas the level in the 7- to 9.5-kHz region was higher by 6 dB in the opposite ear (the levels in the 4.5- to

7-kHz region were equal in the two ears; see Figures 4A and 4B). In one interval of the 2IFC, the higher level in the 2- to 4.5-kHz band favored the left ear, and in the other interval it favored the right ear. As in Part 1, the subjects' task was to determine the order of perceived lateral locations (left–right or right–left). Performance was quantified as the proportion of lateralization judgments consistent with the ILD cue in the 2- to 4.5-kHz region. Each of the 3 subjects completed four runs of 50 trials.

In Part 3, we examined the lateralization of FM pulses that either had identical starting frequencies in the two ears but different ending frequencies or had identical ending frequencies but different starting frequencies. Note that the FM stimuli used in Experiment 1 had different onset and offset frequencies in the two ears (see the Method section and Equations 1 and 3). Stimulus Type 1 started at different frequencies in the two ears (3 and 4 kHz) but terminated at the same frequency (8 kHz), whereas Stimulus Type 2 started at 3 kHz in both ears but terminated at two different frequencies (8 and 9 kHz). The waveforms to the two ears always had equal durations (5 msec) and onset/offset temporal envelopes. These two types of stimuli generated ILD cues either in the 3- to 4-kHz region or in the 8- to 9-kHz region, respectively. The purpose of this part of the experiment was to determine whether an ILD cue at the 3- to 4-kHz region would provide a more salient lateral percept than that at the 8- to 9-kHz region. Note that in Experiment 1, the ITD was carried at both the start and the end of the sweep, hence generating conflicting simultaneous ILD cues. Each subject completed four runs of 50 trials in each of two stimulus conditions. No response feedback was provided at any time during any of the control conditions.

#### **Results and Discussion**

Figure 3 shows the results from Part 1, in which the subjects lateralized an up or down sweep in the absence of response feedback. Each bar shows the averaged data from the 3 subjects. The two left bars represent the data from up sweeps, and the two right bars represent those from down sweeps. The black bars represent the data from trials in which the order of the leading ITDs in the two intervals of the 2IFC were right then left, and the white bars are from trials in which the ordinate shows the percentage of responses that were consistent with lead ITD, with the 50% point representing chance performance. These



Figure 3. Results from Experiment 2A, Part 1. The stimulus was a 1-msec FM pulse that either swept up (left bars) or down (right bars) in frequency. The black bars show the results from trials in which the order of lead ITDs in the two intervals of the trial were right then left, and the white bars show the results from trials in which this stimulus order was left then right. No response feedback was provided. Error bars represent 1 *SD*.



Figure 4. Results from Experiment 2A, Parts 2 and 3. Panels A and B show a sample of noise spectrum used in Part 2. This noise band contained an ILD in the 2- to 4.5-kHz region favoring the right ear and an ILD in the 7- to 9.5-kHz region favoring the left ear. Panel C shows the results from Part 2. Note that the percept of lateralization is dominated by the ILD in the lower frequency region. Panel D shows the results from Part 3. Notice again that the ILD cue in the lower frequency region affects lateralization more than that in the higher frequency region (see the text for details). Error bars represent 1 *SD*.

data clearly verify that a down sweep generates counterintuitive reversals in lateralization that are contrary to that predicted from the stimulus interaural delay and also confirm that the availability of response feedback in Experiment 1 allowed the subjects to adopt a strategy of voting opposite to the side on which they actually perceived the sound.

Figure 4C shows the results for the noise band condition, with each bar showing the data from 1 subject. The ordinate represents the percentage of responses based on the ILD contained in the 2- to 4.5-kHz band (i.e., opposite that of the ILD contained in the 7- to 9.5-kHz band). The data suggest that the ILD contained in the lower frequency region of the spectrum dominates that in the higher region when they are simultaneously present within the same noise band. Figure 4D shows the results from the FM sweep conditions. The black bars represent the data obtained using Stimulus Type 1, with the ILD cue contained in the 3- to 4-kHz region, whereas the white bars show the data obtained using Stimulus Type 2 (ILD in the 8- to 9-kHz region). These data again confirm that the subjects performed significantly better when the ILD cue was contained in the lower frequency region [t(2) = 12.58, p < .01].

## Experiment 2B: Control for Low-Frequency Energy

Short-duration pulses contain substantial low-frequency energy. Since it is well established that the binaural system is extremely sensitive to ITD cues below 1 kHz (Blauert, 1997; Mills, 1960, 1972; Rayleigh, 1907), an important concern was whether ITD cues derived from lowfrequency energy may have contributed to the observed patterns of lateralization and the low thresholds obtained for very short FM pulses (0.5 and 1 msec). To address this issue, we conducted a control experiment using FM sweeps that were high-pass filtered and presented in a background of low-pass masking noise.

## Method

Subjects. Three normal-hearing adults participated in this experiment. Two of the subjects had participated in the earlier experiments, and 1 was a new subject. This latter subject was also experienced in psychoacoustic experiments and practiced the various conditions of the experiment for 1 h prior to data collection.

Apparatus and Procedure. The apparatus and procedure were identical to those described for Experiment 1, with the following exceptions. The stimuli consisted of logarithmic FM up sweeps (3-8 kHz) with durations fixed within a run at 0.5, 1, 5, 10, 25, and 50 msec. We selected only one sweep type and direction, since the results from Experiment 1 suggested that there are no significant differences in performance across sweep types and directions. The stimuli were digitally high-pass filtered in MATLAB using a 512-point FIR (Hamming-windowed) filter with a cutoff frequency of 2 kHz. The output of the digital-to-analog converter was then additionally fed into a dual-channel high-pass analog filter (Kemo BenchMaster VBF 8.13) with a cutoff frequency of 2 kHz and a filter slope of 94 dB/octave. Measurements showed that signal energy was down by 73 dB at 1700 Hz and not detectable at 1500 Hz with our instruments (Brüel & Kjær, Precision Sound Analyzer, Model 2260). The filter outputs were led to the headphones inside the sound booth. In addition, from a separate microprocessor, low-pass-filtered Gaussian noise with a cutoff frequency of 1500 Hz was presented diotically through the headphones at a level of 60 dB SPL. The noise was on continuously throughout the entire run.

#### Results

The results of Experiment 2B are shown in Figure 5. The open symbols show the averaged thresholds from the 3 subjects. For comparison, the averaged data from Experiment 1 are also shown. The results with high-pass filtered stimuli produced results very similar to those with unfiltered sweeps. The slightly better performance with short-duration filtered sweeps relative to that in Experiment 1 is due to the performance of the new subject, who generated thresholds lower than those of the replaced subject at these durations. These findings support our earlier observations and suggest that low-frequency ITD cues below 1 kHz are not responsible for the low thresholds reported in Experiment 1.



Figure 5. Results from Experiment 2B, demonstrating that eliminating low-frequency energy (<1 kHz) did not affect performance. The open circles represent average thresholds from the 3 subjects for FM sweeps high-pass filtered at 2 kHz in a background of continuous low-pass Gaussian noise with a cutoff frequency of 1500 Hz. The results from Figure 1E are replotted for comparison.

## **GENERAL DISCUSSION**

We consider several explanations for the markedly lower ITD thresholds observed for short-duration FM sweeps and the counterintuitive reversals discussed earlier. Prior research with SFM tones suggests that the lateralization of high-frequency FM stimuli is based on an FM-to-AM transformation in the auditory periphery as the instantaneous frequency of the FM sweeps through the passband of cochlear filters (Blauert, 1981; Henning, 1980; Nuetzel & Hafter, 1981; Saberi, 1998; Saberi & Hafter, 1995). Figure 6A shows a 1-msec FM pulse. Figure 6B shows the outputs of a gammatone filter (simulating a cochlear filter; Holdsworth, Nimmo-Smith, Patterson, & Rice, 1988) with a resonant frequency of 5 kHz in response to 1- and 50-msec FM pulses sweeping from 3 to 8 kHz. Note that the filter output is an AM pulse with a steeper slope for the 1-msec sweep than that for the 50-msec sweep. A steeper AM slope provides a more effective ITD envelope cue at high frequencies (Blauert, 1997; Saberi, 1995a, 1998). Figure 6C shows a schematic of a hypothetical dichotic pair of FM sweeps passing through frequency-matched filters centered at 5 kHz. In this diagram, the FM sweep is delayed in the left channel (black), representing an interaural delay favoring the right ear. Gammatone filter outputs for this dichotic stimulus (1-msec sweep from 3 to 8 kHz) are shown in Figure 6D. Note that this stimulus configuration would suggest that the pulse should be lateralized to the right (i.e., perceived intracranially on the right side of the interaural axis).<sup>4</sup> When sweep direction is reversed (Figure 6E) and the ITD favors the left ear, filter outputs shown in Figure 6F predict that the pulse should be lateralized to the left. Surprisingly, both of these stimuli (up and down sweeps) are clearly lateralized to the left, contrary to what would be predicted from a simple FM-to-AM conversion model.

To better understand how the auditory system processes spatial cues from FM-to-AM conversion and why we observe reversals in lateralization, we processed the FM sweeps used in the present experiments through a crosscorrelation model of binaural interaction. This type of computational model, which originated in the theoretical work of Jeffress (1948) and later found neurophysiological support (Carr & Konishi, 1988, 1990; Yin & Chan, 1990), represents an interaural delay as a spatially distributed physiological place map. Peaks of activity along this tonotopically organized frequency  $\times$  delay map correspond to estimated locations in space. Our implementation included preprocessing through a model of the auditory periphery (Figures 7A–7C), consisting of a gammatone filter bank with 50 logarithmically spaced filters from 0.5 to 12 kHz (Holdsworth et al., 1988) and an inner hair-cell model (Meddis, Hewitt, & Shackleton, 1990; Slaney, 1998), followed by cross-correlation of the outputs of corresponding left and right channels with matched filter CFs. The stimuli in this simulation were 1-, 5-, and 10-msec logarithmic up sweeps (similar patterns were observed for down sweeps). The corresponding frequency  $\times$  delay cross-correlation surfaces, for an ITD of 400  $\mu$ sec favoring the right ear (positive lags), are shown in Figures 7D-7F. The arrows



Figure 6. (A) Time waveform of a 1-msec linear FM sweep (3–8 kHz). (B) Outputs of a gammatone filter centered at 5 kHz in response to 1- and 50-msec sweeps (3–8 kHz). (C) Schematic of a dichotic FM up sweep with an ITD favoring the right ear (delayed to the left). (D) Outputs of gammatone filters centered at 5 kHz in response to the stimulus shown in panel C. (E) Schematic of a dichotic FM down sweep with an ITD favoring the left ear. (F) Filter outputs associated with the stimulus shown in panel E.

show the expected positions of straight trajectories of cross-correlation peaks across frequency bands; note the poorly defined main trajectory and spectral splatter associated with the 1-msec FM pulse.

Figures 8A, 8B, 8D, and 8E show the lateralization predictions from this model for the 5-msec sweep. Similar predictions were observed for the 10-msec sweep, as well as for the 1-msec sweep, although in the latter case, the patterns, as was expected, were noisier. Figures 8A and 8B show the predictions for an FM pulse leading to the left ear, and Figures 8D and 8E show those for an FM pulse leading to the right ear. Positive lags denote an ITD favoring the right ear. Figures 8A and 8D show the predictions derived from integrating the cross-correlation activity across all frequency channels (0.5–12 kHz), and Figures 8B and 8E show those from integrating only across low-frequency channels (0.5–1.5 kHz). The latter analysis was included to determine whether spectral energy at low frequencies where carrier interaural delay provides an effective cue to lateralization may have been used. Figures 8A and 8D show that whether predictions of lateral position are based on the lag associated with the peak of the integrated activity or its envelope (Hilbert envelope: dashed curve), an FM sweep leading to the right ear should be perceived on the right, and a sweep leading to the left ear should be perceived on the left. Furthermore, the functions shown in



Figure 7. Panels A–C show the outputs of a computational model of the auditory periphery in response to 1-, 5-, and 10-msec logarithmic FM sweeps, respectively. The model consists of a gammatone filter bank and an inner hair-cell model. Panels D–F show the outputs of a cross-correlation model, with independent inputs from two sets of front-end filter bank–hair-cell computational modules (left/right ears) in response to dichotic FM up sweeps with an ITD of 400  $\mu$ sec favoring the right ear (represented by positive lags in the figures). The arrows show the expected position of straight trajectories of cross-correlation peaks across frequency bands.



Figure 8. (A–D) Predictions of subjective lateral position for a 5-msec logarithmic FM sweep. Panel A shows the predicted position for a down sweep with an ITD favoring the left ear, and panel D shows the results for an up sweep with an ITD favoring the right ear. The black curves show peaks of cross-correlation activity integrated across frequency channels, and the dashed curves show predictions based on the envelopes of cross-correlation activity. Surprisingly, both of these stimuli were perceived on the left. Panels B and E show the same activity patterns generated by integrating only across low-frequency channels (<1.5 kHz). (C) Spectrum of a 5-msec dichotic FM sweep (3–8 kHz) with a 750- $\mu$ sec ITD favoring the right (dashed gray in print; red online) ear. Note an ILD cue favoring the left (black) ear in the 2–4 kHz region marked by the shaded area (see the text). (F) Same as panel C, except that the stimulus was 50 msec in duration.

Figures 8B and 8E suggest that low-frequency cues (from spectral splatter) do not provide sufficient information for predicting lateralization of these pulses, since peak activity (as well as its centroid) occurs at zero lag, predicting an intracranially centered auditory image. This latter prediction is also consistent with the results of Experiment 2B, using high-pass filtered pulses.

As was noted earlier for Experiment 2A, Part 1, both the up and down sweeps shown in Figures 6C and 6E are in fact clearly perceived on the same side of the interaural axis favoring the left-leading channel, contrary to the predictions of cross-correlation analysis. It therefore appears either that ITD cues are not used for the lateralization of directional FM sweeps at high frequencies or that they are dominated by another cue, which we suggest is an ILD cue. Figures 8C and 8F show the spectra of a dichotic 5-msec linear FM sweep (Figure 8C) and a 50-msec sweep (Figure 8F) with the right channel (dashed line) advanced by 750  $\mu$ sec, and hence a start frequency that is higher by approximately 0.75 kHz (note that the waveforms to the left and right channels had identical starting and ending temporal envelopes to eliminate onset/offset ITD envelope cues). Figure 8C shows that this frequency offset produces a strong ILD cue in the 3- to 4-kHz region (shaded area) for the 5-msec sweep but not for the 50-msec sweep. The presence of this ILD cue at 3-4 kHz is independent of sweep direction and consistent with reported percepts of a lateralized image favoring the left ear. This explanation would require that the ILD cue in the 3- to 4-kHz region be more effective than that in the 8- to 9-kHz region, which is in the opposite direction. The results of our control experiments, both with noise bands and FM sweeps, support the idea that the low-frequency ILDs dominate higher frequency ILDs when they are simultaneously present in a complex waveform. The causes of ILD dominance at low frequencies are not relevant to the present study, only that such dominance exists. Possible causes, however, might include an upward spread of masking or asymmetric cortical representation. Widening the bandwidth of the highfrequency ILD region to match, on an octave scale, that of the low-frequency region has no effect on low-frequency dominance, hence ruling out unequal logarithmic widths of the two bands as a cause.

In general, the findings from our model analysis suggest that one should think of listening to high-frequency sounds as potentially dominated by amplitude information, even in situations in which one might think that timing information is the critical cue. Although we do not assert that ITDs are ineffective cues in high-frequency FM waveforms, it would be useful to consider alternative models that incorporate differential weighting of level cues generated from FM stimuli as they sweep through the up or down slopes of an auditory filter and, by extension, the role of ILD cues in the rise/fall slopes of sinusoidally amplitude-modulated (SAM) envelopes. There is considerable psychophysical and neurophysiological evidence suggesting that these types of envelope asymmetries influence both psychophysical performance and the neural encoding of auditory signals (Lu, Liang, & Wang, 2001;

Saberi, 1998; Stecker & Hafter, 2000). There are a variety of conditions under which temporally or spectrally asymmetric neural encoding mechanisms provide a potential ILD cue in either the envelopes of high-frequency SAM sounds or AM envelopes extracted from an FM waveform via band-pass filtering. Determining whether these putative mechanisms are used by the binaural system, however, would require substantial empirical evidence beyond the scope of the present study. Nonetheless, we believe that a more careful consideration of the role of ILDs derived from interaurally time-delayed high-frequency complex waveforms is warranted.

Finally, our findings have broader implications for how natural everyday sounds are processed by the auditory system. FM sweeps form an important component of many communication signals in both humans and animals. Speech sounds contain substantial FM information in the form of rapid frequency glides and formant transitions, which provide cues to phonemic identification (Fant, 1970; Gordon & O'Neill, 1998; A. M. Liberman, Delattre, Gerstman, & Cooper, 1956; Pickett, 1980). Most formant transitions have durations of less than 50 msec, and very brief FM sweeps of 5-20 msec with rates exceeding 30 octaves per second have been used in several studies to examine processing of formant transitions (A. M. Liberman et al., 1956; Luo et al., 2007; Miller & Liberman, 1979). Rapid FM sweeps have also been employed in the study of tonal languages (e.g., Chinese or Thai), where pitch contour variations affect lexical distinction (Howie, 1976; Luo et al., 2007; Stagray, Downs, & Sommers, 1992). The study of FM transitions in speech has also had practical applicability for language processing by children afflicted with language-based learning impairment (LLI). One cause of LLI is speculated to be impaired temporal processing of low-level acoustic features often associated with phonemic processing. Tallal et al. (1996) showed that reducing the FM rate of formant transitions by 50% and amplifying their levels relative to steady-state (non-FM) parts of a phoneme significantly improve the ability of these children to identify speech sounds.

In recent years, the study of frequency sweeps as a common component of communication signals in various animal species has also increased. New World monkey vocalizations (e.g., twitter calls), for example, have been shown to contain FM transitions with rates between 30 and 50 octaves per second (Atencio et al., 2007; Bieser, 1998; Cheung, Bedenbaugh, Nagarajan, & Schreiner, 2001; Nagarajan et al., 2002; Wang, Merzinich, Beitel, & Schreiner, 1995). The role of frequency sweeps in animal communication has also been studied to determine its effect on object identification and informational unmasking (e.g., in nonhuman primates; Egnor, Iguina, & Hauser, 2006; Egnor, Wickelgren, & Hauser, 2007; Petkov, O'Connor, & Sutter, 2003). Some animal communication signals provide mislocalization or poor localization cues that allow conspecific communication without revealing the location of the animal to predators (e.g., cricket chirps or primate alarm calls). These types of sounds are often narrow-band modulated (FM or AM) waveforms.

Other areas of FM research have included auditory scene analysis (Bregman, 1994; Crum & Hafter, 2008), in which the consistency of multiple FM glides in different frequency regions combined with spatial location information can provide powerful cues to auditory object formation and stream segregation in the presence of distracting acoustic information (e.g., other speakers), as well as the study of infant hearing development (Colombo & Horowitz, 1986) and even music perception (d'Alessandro et al., 1998). Findings from the present study contribute further to this growing body of research and to understanding how the auditory system processes this important class of complex sounds.

### AUTHOR NOTE

We thank Ervin R. Hafter, Bruce G. Berg, Michael D'Zmura, and three anonymous reviewers for their helpful comments. This work was supported by NSF Grant BCS0477984 and NIH Grant R01DC009659. Correspondence concerning this article should be addressed to I-Hui Hsieh, Institute of Cognitive Neuroscience, National Central University, Jhongli City, Taoyuan County 32001, Taiwan (e-mail: ihsieh@ncu.edu.tw).

#### REFERENCES

- ATENCIO, C. A., BLAKE, D. T., STRATA, F., CHEUNG, S. W., MERZENICH, M. M., & SCHREINER, C. E. (2007). Frequency-modulation encoding in the primary auditory cortex of the awake owl monkey. *Journal of Neurophysiology*, 98, 2182-2195.
- BAILEY, W. J., GREENFIELD, M. D., & SHELLY, T. E. (1993). Transmission and perception of acoustic signals in the desert clicker, *Ligurotettix coquilletti* (Orthoptera: Acrididae). *Journal of Insect Behavior*, 6, 141-154.
- BERNSTEIN, L. R., & TRAHIOTIS, C. (1982). Detection of interaural delay in high-frequency noise. *Journal of the Acoustical Society of America*, 71, 147-152.
- BIESER, A. (1998). Processing of twitter-call fundamental frequencies in insula and auditory cortex of squirrel monkeys. *Experimental Brain Research*, **122**, 139-148.
- BLAUERT, J. (1981). Lateralization of jittered tones. *Journal of the Acoustical Society of America*, **70**, 694-698.
- BLAUERT, J. (1997). Spatial hearing. Cambridge, MA: MIT Press.
- BREGMAN, A. S. (1994). Auditory scene analysis: The perceptual organization of sound. Cambridge, MA: MIT Press.
- BRILLET, C., & PAILLETTE, M. (1991). Acoustic signals of the nocturnal lizard *Gekko gecko*: Analysis of the long complex sequence. *Bioacoustics*, 3, 33-44.
- BRIMIJOIN, W. O., & O'NEILL, W. E. (2005). On the prediction of sweep rate and directional selectivity for FM sounds from two-tone interactions in the inferior colliculus. *Hearing Research*, **210**, 63-79.
- BUELL, T. N., & HAFTER, E. R. (1988). Discrimination of interaural differences of time in the envelopes of high-frequency signals: Integration times. *Journal of the Acoustical Society of America*, 84, 2063-2066.
- BUTLER, R. A. (1975). The influence of the external and middle ear on auditory discriminations. In W. D. Keidel & W. D. Neff (Eds.), *Hand*book of sensory physiology. Berlin: Springer.
- BUTLER, R. A., & BELENDIUK, K. (1977). Spectral cues utilized in the localization of sound in the median sagittal plane. *Journal of the Acoustical Society of America*, **61**, 1264-1269.
- CARR, C. E., & KONISHI, M. (1988). Axonal delay lines for time measurement in the owl's brainstem. *Proceedings of the National Academy* of Sciences, **85**, 8311-8315.
- CARR, C. E., & KONISHI, M. (1990). A circuit for detection of interaural time differences in the brain stem of the barn owl. *Journal of Neuro*science, **10**, 3227-3246.
- CHEUNG, S. W., BEDENBAUGH, P. H., NAGARAJAN, S. S., & SCHREINER, C. E. (2001). Functional organization of squirrel monkey primary auditory cortex: Responses to pure tones. *Journal of Neurophysiol*ogy, 85, 1732-1749.
- CLOPTON, B. M., WINFIELD, J. A., & FLAMMINO, F. J. (1974). Tonotopic organization: Review and analysis. *Brain Research*, 76, 1-20.

- COLOMBO, J., & HOROWITZ, F. D. (1986). Infants' attentional responses to frequency modulated sweeps. *Child Development*, **57**, 287-291.
- COSCIA, E. M., PHILLIPS, D. P., & FENTRESS, J. C. (1991). Spectral analysis of neonatal wolf vocalizations. *Bioacoustics*, 3, 275-293.
- CROW, G., LANGFORD, T. L., & MOUSHEGIAN, G. (1980). Coding of interaural time differences by some high-frequency neurons of the inferior colliculus: Responses to noise bands and two-tone complexes. *Hearing Research*, 3, 147-153.
- CRUM, P. A., & HAFTER, E. R. (2008). Predicting the path of a changing sound: Velocity tracking and auditory continuity. *Journal of the Acoustical Society of America*, **124**, 1116-1129.
- D'ALESSANDRO, C., ROSSET S., & ROSSI, J. P. (1998). The pitch of shortduration fundamental frequency glissandos. *Journal of the Acoustical Society of America*, **104**, 2339-2348.
- DANKIEWICZ, L. A., HELWEG, D. A., MOORE, P. W., & ZAFRAN, J. M. (2002). Discrimination of amplitude-modulated synthetic echo trains by an echolocating bottlenose dolphin. *Journal of the Acoustical Society of America*, **112**, 1702-1708.
- DEAR, S. P., SIMMONS, J. A., & FRITZ, J. (1993). A possible neuronal basis for representation of acoustic scenes in auditory cortex of the big brown bat. *Nature*, **364**, 620-623.
- EGNOR, S. E. R., IGUINA, C. G., & HAUSER, M. D. (2006). Perturbation of auditory feedback causes systematic perturbation in vocal structure in adult cotton-top tamarins. *Journal of Experimental Biology*, **209**, 3652-3663.
- EGNOR, S. E. R., WICKELGREN, J. G., & HAUSER, M. D. (2007). Tracking silence: Adjusting vocal production to avoid acoustic interference. *Journal of Comparative Physiology*, **193A**, 477-483.
- FANT, G. C. M. (1970). Acoustic theory of speech production. The Hague: Mouton.
- FUZESSERY, Z. M., RICHARDSON, M. D., & COBURN, M. S. (2006). Neural mechanisms underlying selectivity for the rate and direction of frequency-modulated sweeps in the inferior colliculus of the pallid bat. *Journal of Neurophysiology*, **96**, 1320-1336.
- GORDON, M., & O'NEILL, W. E. (1998). Temporal processing across frequency channels by FM selective auditory neurons can account for FM rate selectivity. *Hearing Research*, **122**, 97-108.
- GRANTHAM, D. W. (1984). Interaural intensity discrimination: Insensitivity at 1000 Hz. Journal of the Acoustical Society of America, 75, 1191-1194.
- HAFTER, E. R., & BUELL, T. N. (1990). Restarting the adapted binaural system. *Journal of the Acoustical Society of America*, **88**, 806-812.
- HAFTER, E. R., & DYE, R. H., JR. (1983). Detection of interaural differences of time in trains of high-frequency clicks as a function of interclick interval and number. *Journal of the Acoustical Society of America*, **73**, 644-651.
- HENNING, G. B. (1974). Detectability of interaural delay in highfrequency complex waveforms. *Journal of the Acoustical Society of America*, 55, 84-90.
- HENNING, G. B. (1980). Some observations on the lateralization of complex waveforms. *Journal of the Acoustical Society of America*, 68, 446-454.
- HOLDSWORTH, J., NIMMO-SMITH, I., PATTERSON, R., & RICE, P. (1988). *Implementing a Gammatone filterband*. (SVOS Final Report: Annex C. Part A, the Auditory Filter Bank). Cambridge, U. K.: MRC Applied Psychology Unit.
- HOWIE, J. M. (1976). Acoustical studies of Mandarin vowels and tones. Cambridge: Cambridge University Press.
- HSIEH, I. & SABERI, K. (2007). Temporal integration in absolute identification of musical pitch. *Hearing Research*, 233, 108-116.
- HUBER, F., & THORSON, J. (1985). Cricket auditory communication. Scientific American, 253, 60-68.
- JAVEL, E. (1980). Coding of AM tones in the chinchilla auditory nerve: Implications for the pitch of complex tones. *Journal of the Acoustical Society of America*, 68, 133-146.
- JEFFRESS, L. A. (1948). A place theory of sound localization. Journal of Comparative Physiological Psychology, 41, 35-39.
- KAY, R. H. (1982). Hearing modulation in sounds. *Physiological Review*, 62, 894-975.
- KELLER, C. H., & TAKAHASHI, T. T. (2005). Localization and identification of concurrent sounds in the owl's auditory space map. *Journal of Neuroscience*, 25, 10446-10461.
- KLUMP, G. M., & LANGEMANN, U. (1992). The detection of frequency

and amplitude modulation in the European starling (*Sturnus-vulgaris*): Psychoacoustics and neurophysiology. In Y. Cazals, L. Demany, & K. Horner (Eds.), *Auditory physiology and perception* (pp. 353-359). Oxford: Pergamon.

- LANGNER, G., & SCHREINER, C. E. (1988). Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanisms. *Journal of Neu*rophysiology, **60**, 1799-1822.
- LEVITT, H. L. (1971). Transformed up–down methods in psychophysics. Journal of the Acoustical Society of America, 49, 467-477.
- LIBERMAN, A. M., DELATTRE, P. C., GERSTMAN, L. J., & COOPER, F. S. (1956). Tempo of frequency change as cue for distinguishing classes of speech sounds. *Journal of Experimental Psychology*, **52**, 127-137.
- LIBERMAN, M. C. (1982). The cochlear frequency map for the cat: Labeling auditory nerve fibers of known characteristic frequency. *Journal* of the Acoustical Society of America, **72**, 1441-1449.
- LU, T., LIANG, L., & WANG, X. Q. (2001). Neural representations of temporally asymmetric stimuli in the auditory cortex of awake primates. *Journal of Neurophysiology*, 85, 2364-2380.
- LUO, H., BOEMIO, A., GORDON, M., & POEPPEL, D. (2007). The perception of FM sweeps by Chinese and English listeners. *Hearing Re*search, 224, 75-83.
- MCFADDEN, D., & MOFFITT, C. M. (1977). Acoustic integration for lateralization at high frequencies. *Journal of the Acoustical Society of America*, 61, 1604-1608.
- MCFADDEN, D., & PASANEN, E. G. (1975). Binaural beats at high frequencies. Science, 190, 394-396.
- McFADDEN, D., & PASANEN, E. G. (1976). Lateralization at high frequencies based on interaural time differences. *Journal of the Acoustical Society of America*, **59**, 634-639.
- MEDDIS, R., HEWITT, M. J., & SHACKLETON, T. M. (1990). Implementation details of a computation model of the inner hair-cell/auditorynerve synapse. *Journal of the Acoustical Society of America*, 87, 1813-1816.
- MENDELSON, J. R., & CYNADER, M. S. (1985). Sensitivity of cat primary auditory cortex (AI) neurons to the direction and rate of frequency modulation. *Brain Research*, 327, 331-335.
- MENDELSON, J. R., & GRASSE, K. L. (1992). A comparison of monaural and binaural responses to frequency modulated (FM) sweeps in cat primary auditory cortex. *Experimental Brain Research*, 91, 435-454.
- MENDELSON, J. R., SCHREINER, C. E., SUTTER, M. L., & GRASSE, K. L. (1993). Functional topography of cat primary auditory-cortex responses to frequency-modulated sweeps. *Experimental Brain Re*search, 94, 65-87.
- MILLER, J. L., & LIBERMAN, A. M. (1979). Some effects of lateroccurring information on the perception of stop consonant and semivowel. *Perception & Psychophysics*, 25, 457-465.
- MILLS, A. W. (1960). Lateralization of high-frequency tones. Journal of the Acoustical Society of America, 32, 132-134.
- MILLS, A. W. (1972). Auditory localization. In J. V. Tobias (Ed.), Foundations of modern auditory theory (Vol. 2, pp. 301-345). New York: Academic Press.
- MOORE, B. C. J., & SEK, A. (1992). Detection of combined frequency and amplitude modulation. *Journal of the Acoustical Society of America*, **92**, 3119-3131.
- MÜLLER, M., VON HÜNERBEIN, K., HOLDIS, S., & SMOLDERS, J. W. T. (2005). A physiological place-frequency map of the cochlea in the CBA/J mouse. *Hearing Research*, **202**, 63-73.
- NAGARAJAN, S. S., CHEUNG, S. W., BEDENBAUGH, P., BEITEL, R. E., SCHREINER, C. E., & MERZENICH, M. M. (2002). Representation of spectral and temporal envelope of twitter vocalizations in common marmoset primary auditory cortex. *Journal of Neurophysiology*, 87, 1723-1737.
- NELKEN, I., & VERSNEL, H. (2000). Responses to linear and logarithmic frequency-modulated sweeps in ferret primary auditory cortex. *European Journal of Neuroscience*, **12**, 549-562.
- NUETZEL, J. M., & HAFTER, E. R. (1976). Lateralization of complex waveforms: Effects of fine structure, amplitude and duration. *Journal* of the Acoustical Society of America, **60**, 1339-1346.
- NUETZEL, J. M., & HAFTER, E. R. (1981). Lateralization of complex waveforms: Spectral effects. *Journal of the Acoustical Society of America*, 69, 1112-1118.
- PETKOV, C. I., O'CONNOR, K. N., & SUTTER, M. L. (2003). Illusory

sound perception in macaque monkeys. *Journal of Neuroscience*, 23, 9155-9161.

- PICKETT, J. M. (1980). The sounds of speech communication: A primer of acoustic phonetics and speech perception. Baltimore: University Park Press.
- RAYLEIGH, L. (1907). On our perception of sound direction. *Philosophi*cal Magazine, 13, 214-232.
- RAZAK, K. A., & FUZESSERY, Z. M. (2006). Neural mechanisms underlying selectivity for the rate and direction of frequency-modulated sweeps in the auditory cortex of the pallid bat. *Journal of Neurophysiology*, **96**, 1303-1319.
- RAZAK, K. A., & FUZESSERY, Z. M. (2007). Development of inhibitory mechanisms underlying selectivity for the rate and direction of frequency-modulated sweeps in the auditory cortex. *Journal of Neuroscience*, 27, 1769-1781.
- ROBISSON, P., AUBIN, T., & BREMOND, J.-C. (1993). Individuality in the voice of the emperor penguin *Aptenodytes forsteri*: Adaptation to a noisy environment. *Ethology*, 94, 279-290.
- ROMANI, G. L., WILLIAMSON, S. J., & KAUFMAN, L. (1982). Tonotopic organization of the human auditory cortex. *Science*, 216, 1339-1340.
- RYAN, M. J., & WILCZYNSKIN, W. (1988). Coevolution of sender and receiver: Effect on local mate preference in cricket frogs. *Science*, 240, 1786-1788.
- SABERI, K. (1995a). Lateralization of comodulated complex waveforms. Journal of the Acoustical Society of America, 98, 3146-3156.
- SABERI, K. (1995b). Some considerations on the use of adaptive methods for measuring interaural-delay thresholds. *Journal of the Acoustical Society of America*, 98, 1803-1806.
- SABERI, K. (1998). Modeling interaural delay sensitivity to frequency modulation at high frequencies. *Journal of the Acoustical Society of America*, 103, 2551-2564.
- SABERI, K., & HAFTER, E. R. (1995). A common neural code for frequency- and amplitude-modulated sounds. *Nature*, 374, 537-539.
- SABERI, K., & PERROTT, D. R. (1999). Cognitive restoration of reversed speech. *Nature*, **398**, 760.
- SABERI, K., & PETROSYAN, A. (2004). A detection-theoretic model of echo inhibition. *Psychological Review*, **111**, 52-66.
- SABOURIN, P., GOTTLIEB, H., & POLLACK, G. S. (2008). Carrierdependent temporal processing in an auditory interneuron. *Journal of* the Acoustical Society of America, **123**, 2910-2917.
- SIMMONS, J. A. (1979). Perception of echo phase information in bat sonar. Science, 204, 1336-1338.
- SLANEY, M. (1998). Auditory toolbox: A MATLAB toolbox for auditory modeling work (Tech. Rep. 1998-010). Palo Alto, CA: Interval Research Corporation.
- SMITH, R. L., & BRACHMAN, M. L. (1980). Response modulation of auditory nerve fibers by AM stimuli: Effects of average intensity. *Hearing Research*, 2, 123-133.
- STAGRAY, J. R., DOWNS, D., & SOMMERS, R. K. (1992). Contributions of the fundamental, resolved harmonics, and unresolved harmonics in tone–phoneme identification. *Journal of Speech & Hearing Research*, 35, 1406-1409.
- STECKER, G. C., & HAFTER, E. R. (2000). An effect of temporal asymmetry on loudness. *Journal of the Acoustical Society of America*, 107, 3358-3368.
- TALLAL, P., MILLER, S. L., BEDI, G., BYMA, G., WANG, X., NAGARAJAN, S. S., ET AL. (1996). Language comprehension in language-learning impaired children improved with acoustically modified speech. *Sci*ence, 271, 81-84.
- WANG, X., MERZENICH, M. M., BEITEL, R., & SCHREINER, C. E. (1995). Representation of a species-specific vocalization in the primary auditory cortex of the common marmoset: Temporal and spectral characteristics. *Journal of Neurophysiology*, 74, 2685-2706.
- WETHERILL, G. B., & LEVITT, H. (1965). Sequential estimation of points on a psychometric function. *British Journal of Mathematical Psychol*ogy, 18, 1-10.
- WIGHTMAN, F. L., & KISTLER, D. J. (1989a). Headphone simulation of free-field listening: I. Stimulus synthesis. *Journal of the Acoustical Society of America*, 85, 858-867.
- WIGHTMAN, F. L., & KISTLER, D. J. (1989b). Headphone simulation of free-field listening: II. Psychophysical validation. *Journal of the Acoustical Society of America*, 85, 868-878.

YIN, T. C. T., & CHAN, J. C. K. (1990). Interaural time sensitivity in medial superior olive of cat. *Journal of Neurophysiology*, 64, 465-488.

- Yost, W. A. (1981). Lateral position of sinusoids presented with interaural intensive and temporal differences. *Journal of the Acoustical Society of America*, **70**, 397-409.
- YOST, W. A., & DYE, R. H., JR. (1988). Discrimination of interaural differences of level as a function of frequency. *Journal of the Acoustical Society of America*, 83, 1846-1851.
- Yost, W. A., & HAFTER, E. R. (1987). Lateralization. In W. A. Yost & G. Gourevitch (Eds.), *Directional hearing* (pp. 49-84). New York: Springer.
- ZHOU, B., GREEN, D. M., & MIDDLEBROOKS, J. C. (1992). Characterization of external ear impulse responses using Golay codes. *Journal of* the Acoustical Society of America, 92, 1169-1171.

#### NOTES

1. All prior studies of the lateralization of FM sounds have used sinusoidal FMs that have bandwidths of from a few hertz to, at most, 0.5 kHz, whereas the directional FMs used here sweep through wide regions of the spectrum ( $\geq$ 5 kHz).

2. A 0.5- or 1-msec pulse is effectively a complex broadband click. However, we have found that an up-frequency 1-msec pulse is distinguishable from a down-frequency pulse, suggesting that the temporal asymmetry of the pulse is encoded by the system. Note also that given the high starting frequency of the sweep, several full periods of the waveform are contained within these brief pulses (see Figures 6A–6F). Hsieh and Saberi (2007) have shown that very brief tones of a few cycles may generate an identifiable pitch.

3. For a fixed frequency range, the rate and duration of a sweep are, of course, inversely related, and their effects cannot be independently determined. To determine the effects of sweep rate separate from duration would require changing the sweep's frequency range. For the present study, we decided to keep sweep range constant and to focus on the effects of duration/rate on ITD sensitivity.

4. Several studies have shown that single high-pass filtered clicks can easily be lateralized on the basis of an ITD cue, with most thresholds averaging under 100  $\mu$ sec (Buell & Hafter, 1988; Hafter & Buell, 1990; Hafter & Dye, 1983; Saberi & Petrosyan, 2004). Bernstein and Trahiotis (1982) emphasized the importance of eliminating low-frequency energy (<1 kHz) when measuring ITD thresholds for high-pass filtered sounds. They showed that eliminating residual low-frequency energy in high-pass noise (i.e., by increasing the high-pass filter slope) substantially elevates ITD thresholds. Nonetheless, even with elimination of low-frequency energy, they show that subjects are able to lateralize high-pass (>3-kHz) noise bands, with thresholds ranging from 70 to 100  $\mu$ sec at a 75% correct-detection level.

(Manuscript received June 26, 2008; revision accepted for publication June 25, 2009.)