Directional hearing is only weakly dependent on the rise time of acoustic stimuli

Rüdiger Krahe

Institute of Biology, Department of Behavioral Physiology, Humboldt University, Invalidenstr. 43, D-10115 Berlin, Germany

Ole Naasbye Larsen

Centre for Sound Communication, Institute of Biology, Odense University, Campusvej 55, DK-5230 Odense M, Denmark

Bernhard Ronacher

Institute of Biology, Department of Behavioral Physiology, Humboldt University, Invalidenstr. 43, D-10115 Berlin, Germany

(Received 14 May 1999; revised 30 September 1999; accepted 25 October 1999)

First-spike latency differences between left and right auditory-nerve fibers have been proposed as one of the physiological cues for sound localization. Since first-spike latency depends not only on stimulus intensity, but also on the steepness of the amplitude rise of a sound stimulus, differences in first-spike latency are not a simple function of interaural level differences but also a function of stimulus rise time. We therefore investigated whether rise time influences human directional hearing in a localization paradigm. Subjects tended to localize a 3-kHz tone pulse with a long (18-ms) rise time further to the side than one with a short (2-ms) rise time delivered from the same source. The small size of this effect and its large inter-individual variability, however, suggest that it is of minor importance for human directional hearing. © 2000 Acoustical Society of America.

PACS numbers: 43.66.Qp, 43.66.Pn, 43.66.Mk, 43.66.Ba [DWG]

INTRODUCTION

Sound direction may be inferred from monaural spectral cues and from a comparison of the inputs to the two ears. The physical cues available for this comparison are interaural differences in level (interaural level differences: ILDs) and in time (interaural time delays: ITDs). The difference in the timing of neuronal activity between left and right auditory pathway induced by the physical ITD may be enlarged by the ILD, due to the dependence of first-spike latency on stimulus intensity (Kitzes et al., 1978; Heil and Irvine, 1997). These ILD-dependent differences in latency have been proposed to be the basis for ILD processing in the central nervous system (Yin et al., 1985). Differences in the timing of first spikes may also be important for the precedence effect (or law of the first wavefront; Wallach et al., 1949), i.e., the observation that in reverberant environments the first incoming wavefront of a sound stimulus is the most salient cue for directional hearing. However, there is strong physiological evidence that first-spike temporal differences do not only depend on ILD (and, of course, on ITD) but also on the rise time of the respective stimulus. The first-spike latency of auditory-nerve (AN) fibers is determined by the steepness of the envelope rise which in turn depends on rise time and peak level (in case of a linear rise function; Heil and Irvine, 1996, 1997). For shallow slopes (long rise time and/or low intensity) latency-intensity functions steeply decrease to approach the minimum latency of the respective cell asymptotically (Fig. 1). The shape of this curve has an important implication: For a given ILD and stimulus intensity, a long-rise-time stimulus will lead to a large latency difference (ΔT in Fig. 1) between the AN fibers of the two ears while latency differences will be small for short-rise-time stimuli. Therefore, if our central nervous system heavily relies on first-spike latency differences for assessing the angle of sound incidence, directional hearing may be influenced by the rise time of the respective stimulus. Due to larger latency differences, a long-rise-time stimulus could convey a more lateralized percept of the sound than a short-rise-time stimulus delivered from the same source.

To test this hypothesis we conducted a localization experiment under free-sound-field conditions. Subjects had to decide whether the second stimulus of a pair of sequentially presented stimuli was delivered from a more frontal or a more lateral position than the first stimulus. We predicted that subjects localize long-rise-time stimuli further to the side than short-rise-time stimuli, due to the larger latency differences induced by the long rise time.

I. METHODS

The free-field experiment was performed in the anechoic room at Odense University and was approved by the local ethics committee according to Danish guidelines. The three authors and nine inexperienced subjects from the Institute of Biology at Odense University volunteered for the experiments (six females, age range 24–41 years; six males, age range 24–49 years). All subjects were checked for normal hearing with a standard audiometer (Bosch type KS 5).
The stimuli were 20-ms tone pulses with carrier frequencies of 0.75 kHz, 3 kHz, or 5 kHz, respectively. The tone pulses were linearly amplitude modulated with a rise time of 18 ms directly followed by a decay of 2 ms or with a rise time of 2 ms directly followed by a decay of 18 ms [short-rise-time stimuli; see insets in Fig. 2(a)]. Each trial consisted of two stimuli of the same carrier frequency separated by a pause of 1 s. The first stimulus of each trial—the reference stimulus—was always delivered from a loudspeaker situated at 30° to the left of the subject, the second from one of seven speakers, chosen in a pseudorandom order. Speaker positions were 10°, 22°, 26°, 30°, 34°, 38°, and 50° to the left of the subject. This speaker arrangement was chosen after preliminary tests had shown that the task was too easy when the reference speaker was positioned at 0°. All possible stimulus combinations were employed: (1) both stimuli with long rise time, (2) both with short rise time, (3) first long, second short, (4) first short, second long [see Fig. 2(a)]. The subjects had to decide whether the percept of the second sound source was located frontal or lateral relative to the first one. Thus the experiment was a relative localization task, not an absolute one. One session consisted of 56 trials (four stimulus combinations, seven speakers, two presentations, and one carrier frequency). Most subjects attended two sessions per day. Data for a given carrier frequency are based on eight to ten sessions per subject. Five subjects participated in the 0.75-kHz condition, 12 subjects in the 3-kHz, and 6 subjects in the 5-kHz conditions.

During an experiment a subject was seated in a chair in the anechoic room, the ceiling, walls, and floor of which were covered with 55-cm-long mineral wool wedges (room dimensions: 2.3×2.7×3.5 m between wedge tips; room anechoic above approx. 200 Hz; floor made without wire mesh). The subject’s head rested comfortably on a chin holder. The seven speakers were located in a horizontal semicircle 2.58 m from the middle of the subject’s head with their diaphragm centers at the height of the subject’s ears. Subjects were instructed to fixate on a cross on the wall and keep their head as immobile as possible during a session. After each trial, listeners reported their decision (“frontal” or “lateral”) via an intercom to the control room. No information on the correctness of the decision was given.

The stimulus waveforms had been calculated and stored digitally in a personal computer. Following D/A conversion by a two-channel 12-bit DT2821 board (DATA TRANSLATION) at a 40-kHz sampling rate, the signals were fed through a manual attenuator (HATFIELD INSTRUMENTS, type 2125), a custom-built reconstruction filter (6-kHz low pass), and a stereo power amplifier (DENON type P0A-2400). A custom-built digital switch routed the signal to the chosen speaker. The speakers (JBL, type Control 1G, two-way speakers) were selected such that the sound intensity varied as little as possible at the position occupied by the subject’s head (variation from speaker to speaker was 1.9–
3.6 dB as measured with a half inch microphone, Brüel & Kjaer type 4133, connected to a measuring amplifier, B&K type 2636). Stimulus intensity during the experimental sessions was set to 45 dB SPL (peak), measured without a subject being present.

For each subject we calculated the percentage of decisions “frontal” as a function of the angle between the reference sound source 1 and sound source 2, with each percentage value based on 16 to 20 decisions. The percentage data were arcsine transformed and linear regressions were calculated. In order to exclude the saturation ranges, only the values between 10% and 90% and the first value equal to or below 10% and equal to or above 90%, respectively, were included. To measure the influence of the envelope cues on localization we first determined the intersection of the regression line with the horizontal for 50% decisions “frontal” for both long–short and short–long stimulus combinations and then calculated the angular difference between these two intersections [cf. Fig. 2(b)–(d)].

II. RESULTS

The subjects reported whether they perceived the second stimulus from a more frontal or a more lateral position than the first stimulus. The probability of the decision “frontal” is shown in Fig. 2(a) for subject AMS as a function of the angle between the source of the second tone pulse and the reference speaker (at 30° to the left) for the four different stimulus combinations. The prediction was that the psychometric curves would be shifted to more lateral positions when a long-rise-time stimulus was referenced to a short-rise-time stimulus and vice versa. The shift was calculated from the intersections of the linear regressions with the 50% horizontal [Fig. 2(a)]. These angular differences are shown for all subjects in the histograms in Fig. 2(b), (c), and (d). Values close to 0° indicate that the regressions for the long–short and short–long condition were virtually identical for the respective subjects. At 0.75 kHz and at 5 kHz there was no obvious deviation from a distribution around 0° [Fig. 2(b) and (d); \( p = 0.9 \) for 0.75 kHz; \( n = 5 \); \( p = 0.6 \) for 5 kHz; \( n = 6 \); Wilcoxon’s matched pairs signed rank test (Sachs, 1992)]. At 3 kHz we observed a statistically significant shift of the distribution to positive values (\( p = 0.05, n = 12 \)). Thus at least for 3 kHz, the prediction that long-rise-time stimuli are localized further to the side than stimuli with a short rise time was upheld. The median of the angular difference was 3.7°. This rather moderate median value and the occurrence of negative values in a quarter of the subjects tested, however, indicate that the effect of rise time on directional hearing in the free sound field is not very robust [Fig. 2(c)].

In most cases (17/23) the psychometric curves for the control trials (long–long and short–short) fell between the curves for the test stimuli (long–short and short–long), indicating that a lateral shift of the test curves was not simply due to a long or a short rise being presented as the reference stimulus. As one would expect for a relative localization task, we found no systematic shift of the curves for long–long relative to those for short–short (\( p > 0.1; n = 12 \) for 3 kHz). Neither did the 50% intersections for long–long or short–short deviate from 0° (\( p > 0.1 \)).

In order to test whether stimulus combination affected the steepness of the psychometric curves [Fig. 2(a)], we compared the slopes of the respective regression lines. The slope was calculated from the angular difference between the 25%- and 75%-intersection points of the regression lines. This angular-difference threshold thus is a measure analogous to the minimum audible angle described by Mills (1958). The slopes for the long–long, short–short, long–short, and short–long conditions did not differ significantly according to a Kruskal–Wallis test (\( H = 2.086, 1.128, \) and 2.21 for 0.75 kHz, 3 kHz, and 5 kHz, respectively; \( df = 3 \) in each case).

Carrier frequency, however, significantly affected the steepness of the psychometric curves. The angular-difference thresholds for 0.75 kHz were only half as large as the thresholds for both 3 kHz and 5 kHz (\( p < 0.001 \) and \( p < 0.01 \), respectively; Kruskal–Wallis test), whereas angular-difference thresholds did not differ significantly between 3 kHz and 5 kHz. This conforms well with the finding that at frequencies below 1.5 kHz ITDs in the stimulus fine structure are the most salient directional cue, overriding possibly conflicting envelope delays (Henning, 1983; see also Zurek, 1993).

III. DISCUSSION

Our study was aimed at testing the hypothesis that long rise times of sound pulses entail stronger directional cues than very short ones presumably due to larger first-spike latency differences between left and right AN fibers (Fig. 1). We found a statistically significant influence of rise time only at 3 kHz. The effect pointed in the predicted direction [Fig. 2(c)]. However, the small absolute size of the effect and the large inter-individual variance indicate that rise time is not a robust determinant of directional information, at least under our experimental conditions.

This outcome is somewhat surprising since interaural disparities at stimulus onset are important indicators of the direction to a sound source, especially in reverberant environments (Wallach et al., 1949; Zurek, 1987), and it seems plausible that first-spike latency differences are involved in this analysis. According to physiological data, these latency differences depend on stimulus rise time (Heil and Irvine, 1997). At least at 0.75 kHz, the negative result of our experiment [Fig. 2(b)] may be due to ringing of the auditory filters. It is reasonable to assume that ringing reduces the actual differences between the long- and short-rise-time stimulus envelopes at the level of the cochlea. However, since ringing effects increase with decreasing center frequency and the concurrent decreasing bandwidth of auditory filters (de Boer and Kruidenier, 1990; Evans, 1989) this explanation cannot account for the effects of rise time at 3 kHz. We cannot exclude that a much larger sample size or well-trained listeners might have rendered statistically significant results (note that the decisions of the three authors did not differ from those of the nine inexperienced subjects). Our interest, however, was in the existence of a behaviorally relevant effect of rise time upon sound localization. This can be excluded as
the effects found are still in the range of the minimum audible angle as determined by Mills (1958; see also Hartmann, 1983).

Since the physiological evidence for the dependence of first-spike latency differences on rise time is strong (cat: Heil and Irvine, 1997; for comparable work on insects: Adam, 1977; Krahe and Ronacher, 1993), it is conceivable that directional hearing relies more heavily on other neuronal cues. An ultimate reason for this may be that a strong dependence of directional hearing on latency differences would entail the risk of localization errors since the latency difference not only depends on stimulus direction (ILD) but also on the shape of stimulus onset and on intensity. In this context it is interesting to note that owls, which rely heavily on acoustic cues for locating prey, do not use onset transients for sound localization but rather focus on ongoing ITDs (Moisseff and Konishi, 1981). In a behavioral paradigm, Martin and Webster (1987) also found evidence that cats do not use transient temporal disparities to a significant amount. However, since humans undoubtedly can use temporal onset disparities (Wallach et al., 1949; Zurek, 1987), we tentatively suggest that our nervous system rates other onset cues higher than just the first-spike temporal difference arising between left and right auditory pathway.

ACKNOWLEDGMENTS

This work would have been impossible without the collaboration of members of the Odense laboratory who volunteered as experimental subjects. We gratefully acknowledge the very helpful comments of Drs. Peter Heil, Walter Metzner, the anonymous referees, and the editor at various stages of the manuscript. The Centre for Sound Communication is financed by the Danish National Research Foundation. This research was supported by the Deutsche Forschungsgemeinschaft (grant to B.R.).


