Hearing Research 312 (2014) 103–113

Contents lists available at ScienceDirect

Hearing Research

journal homepage: www.elsevier.com/locate/heares

Research paper

Modeling the utility of binaural cues for underwater sound localization

Jennifer N. Schneider*, David R. Lloyd, Patchouly N. Banks, Eduardo Mercado III

Department of Psychology, University at Buffalo, The State University of New York Park Hall, Room 204, Buffalo, NY 14260, USA

ARTICLE INFO

Article history:
Received 25 June 2013
Received in revised form
10 March 2014
Accepted 24 March 2014
Available online 13 April 2014

ABSTRACT

The binaural cues used by terrestrial animals for sound localization in azimuth may not always suffice for accurate sound localization underwater. The purpose of this research was to examine the theoretical limits of interaural timing and level differences available underwater using computational and physical models. A paired-hydrophone system was used to record sounds transmitted underwater and recordings were analyzed using neural networks calibrated to reflect the auditory capabilities of terrestrial mammals. Estimates of source direction based on temporal differences were most accurate for frequencies between 0.5 and 1.75 kHz, with greater resolution toward the midline (2°), and lower resolution toward the periphery (9°). Level cues also changed systematically with source azimuth, even at lower frequencies than expected from theoretical calculations, suggesting that binaural mechanical coupling (e.g., through bone conduction) might, in principle, facilitate underwater sound localization. Overall, the relatively limited ability of the model to estimate source position using temporal and level difference cues underwater suggests that animals such as whales may use additional cues to accurately localize conspecifics and predators at long distances.

1. Introduction

Vertebrates vary greatly in their ability to localize sounds, as well as in the mechanisms that enable them to do so (Fay and Popper, 2005). Aquatic animals are especially versatile (Fay and Popper, 2005; Mooney et al., 2012). Comparisons of hearing abilities across diverse species can provide important clues about how basic auditory mechanisms function and how those mechanisms originally evolved (Manley and Fuchs, 2011). For example, most current theories of the neural mechanisms of sound localization in humans were developed based on studies of non-humans (King and Middlebrooks, 2011; Phillips et al., 2012). Sound localization mechanisms are thought to be comparable across most terrestrial mammals (Phillips et al., 2012), although not all mechanisms appear to be used by all species (Heffner and Heffner, 1992a) and sound localization ability varies. Here, we assess whether current models of sound localization by terrestrial mammals can provide insights into how mammals localize sound sources underwater.

Mammals typically use two main binaural differences to localize sound sources in azimuth (i.e., in the horizontal plane; for review, see Brown and May, 2005; Heffner and Heffner, 1992a): differences in time of sound arrival (interaural temporal differences; ITDs) and differences in received sound pressure level (interaural level differences; ILDs). Some species may also use spectral cues produced by head and body reflections (e.g., Aytekin et al., 2004), although these cues are usually more critical for estimating the elevation of a sound source. All of these cues become less reliable indicators of source location underwater. The difficulties arise because sound travels five times faster in water than in air and can easily pass from water into tissues and bones within the head (e.g., Bailey et al., 2013; Savel et al., 2009). Theoretically, these transformations should greatly decrease both ITDs and ILDs, and should eradicate normal spectral cues arising from head and body surfaces. For example, harbor seals localize high frequency sounds (over 2 kHz) better in air than in water (Bodson et al., 2007). Nevertheless, humans retain the ability to localize sounds underwater (Feinstein, 1973; Norman et al., 1971; Savel et al., 2009; Shipley et al., 1991; Wilhelm et al., 2011). Additionally, species such as dolphins that

---

Abbreviations: ANOVA, analysis of variance; AHRS, altitude heading reference system; AM, amplitude modulated; BRS, biomimetic recording system; CFP, cochlear frequency position; HRTF, head related transfer function; ILD, interaural level difference; ITD, interaural timing difference; MAA, minimum audible angle; PVC, polyvinyl chloride

* Corresponding author. Present address: Department of Social Sciences, LCC International University, Kretingos 36, Klaipeda LT-92307, Lithuania. Tel.: +37 046310745.

E-mail address: jns5@buffalo.edu (J.N. Schneider).

http://dx.doi.org/10.1016/j.heares.2014.03.011
0378-5955/

are adapted to hearing sounds underwater can, in some cases, localize sounds underwater better than many terrestrial mammals can localize sound in air (Au, 1993).

Some aquatic animals (e.g., fish) are known to use mechanisms for localizing sound sources underwater that differ in many respects from those used by terrestrial mammals (Fay, 2005). In contrast, whales and dolphins are generally assumed to use the same ITD and ILD cues as used by terrestrial mammals (Mooney et al., 2012). Preliminary behavioral evidence from manatees suggest that these animals use ITDs, but not ILDs (Colbert et al., 2009). Dolphins, in addition to using binaural cues, may use monaural spectral cues (specified by internal head related transfer functions, HRTFs), produced by their lower jaw and fat channels, to localize high-frequency sound sources either passively or actively (Aroyan, 2001; Branstetter et al., 2007; Ketten, 2000; Supin and Popov, 1993). Recently, researchers discovered a distinct type of fatty tissue near the ears of a baleen whale that could potentially function as a similar sound reception channel and a potential source of monaural cues (Yamato et al., 2012).

In the current study, we developed a computational model to predict the source localization cues available to mammals using ITDs and ILDs. The model instantiates basic principles of sound localization in terrestrial mammals and is comparable to earlier neural network models developed to simulate sound localization by dolphins (Branstetter et al., 2007) and humans (Backman and Karljalainen, 1993; Chung et al., 2000; Datum et al., 1996; Janko et al., 1995; Jin et al., 2000; Neti et al., 1992; Palmieri et al., 1991; Schauer et al., 2000). First, the model was tested using anatomical data from several mammals (elephants, cattle, cats, and humans) to establish that it was capable of predicting key aspects of the known azimuthal resolution in these animals. The model was then used to simulate underwater sound localization in humans and humpback whales. Although the comparative results do not guarantee that the model can be generalized to humpback whales, we believe that a clear demonstration of the model's ability to predict the performance of terrestrial mammals gives credence to the plausibility of the model as a predictor of spatial hearing in whales, assuming that whales utilize the same localization cues available to terrestrial mammals. Finally, the model was tested on its ability to determine the azimuth of a source broadcasting various sounds underwater. We hypothesized that both ITD and ILD cues would provide information useful for localizing the types of sounds heard by humpback whales.

2. Simulations of airborne sound localization by terrestrial mammals

The goal of this initial analysis was to develop, calibrate, and evaluate a neural network model of sound localization in azimuth using ITDs and ILDs. Neural network models were developed for elephants as representatives of large terrestrial mammals, cattle (which have interaural distances similar to whales), and cats (which have a head size to wavelength ratio in air similar to that of whales in water). Physiological and anatomical measures used to determine parameter settings in these simulations (Heffner and Heffner, 1982, 1992b; LePage, 2003; West, 1985) are reported in Table 1. Fig. 1 provides a conceptual overview of how the neural network model estimates sound source azimuth (a more computationally detailed depiction is provided in the Supplementary Materials).

2.1. Methods

2.1.1. ITDs in air

ITDs reflect differences in sound onset, modulations of the amplitude envelope, and phase differences in the fine structure of the signal (Brown and May, 2005; Heffner and Heffner, 1992a). In air, ITD can be estimated using the following equation:

$$\mathrm{ITD} = \frac{r}{c}(\theta + \sin\theta) \tag{1}$$

where c is the speed of sound, r is the radius of the head, and θ is the angle between the center line between the two sensors and the sound source in radians (Brown and May, 2005).

The utility of phase differences for localizing sound sources is frequency dependent; specifically, there are upper and lower bounds on what frequencies an animal can use to detect ITDs in pure tones. These boundaries are determined partially by the ability of neurons in the periphery to phase lock to low and high frequency sound (Heffner and Heffner, 1992a; Rose et al., 1967). At high frequencies, the use of ITDs is also limited by the length of the wavelength relative to the interaural distance (wavelengths less than or equal to twice the interaural distance correspond to source locations on the left and right). At low frequencies, the phase differences between the two ears can become so small that they are not detectable. Behavioral data in humans indicate discrimination of ITDs as small as 10–20 μs in specific conditions (Dye and Hafter, 1984; Mills, 1958), while single neuron responses in the inferior colliculus of guinea pigs show discrimination abilities of 20–30 μs (Skottun et al., 2001). While ITDs are greatest at the periphery, changes in ITDs are smaller per angular degree at the periphery than near the midline. Therefore, the smallest angle of resolution in which an animal can detect differences in source location, called the minimum audible angle (MAA), tends to be greatest when sounds are coming from near the midline.

Sixty dual channel synthetic pure tones between 0.05 and 3.00 kHz (44.1 kHz sampling rate, 16 bits, 1.0 s, rectangular amplitude envelope) were used to assess model performance. Simulated ITDs were created by offsetting one channel from the other for each integer angle between −90 and 90°; offsets were calculated using the following equation:

$$\mathrm{offset} = \mathrm{round}\left(\frac{\sin\theta^{*}d^{*}f_s}{c}\right) \tag{2}$$

Table 1
Head and basilar membrane properties used in calculating timing and level differences based on anatomical and behavioral measurements and estimates of mammals used in this study.[a]

Mammal	Interaural distance (m)	Head radius assuming spherical shape (m)	Basilar membrane length (mm)	Max. Frequency threshold at 60 dB SPL (kHz)	Species specific constant k	Level difference discrimination threshold (Hz)[b]
Elephant	1.15	0.37	60	10	1.0	50
Cow	0.44	0.14	38	35	1.0	150
Cat	0.09	0.03	23	77	0.8	600
Human	0.10	0.09	34	19	1.0	850
Whale	0.40	1.00	58	30	1.0	100

[a] Non-whales: Heffner and Heffner, 1982; Heffner and Heffner, 1992a,b; West, 1985; LePage, 2003. Whales: Ketten, 1994; Ketten 1997; Mercado, 1998.

[b] Frequency where wavelength is 1/10 the size of the diameter of the head.