Psychoacoustic investigation of auditory cues involved in human underwater sound localization

Sophie Savel, Carolyn Drake

To cite this version:

Sophie Savel, Carolyn Drake. Psychoacoustic investigation of auditory cues involved in human underwater sound localization. 5th European Conference on underwater acoustics, Jul 2000, Villeurbane, France. 2000. <hal-01128128>

HAL Id: hal-01128128
https://hal.archives-ouvertes.fr/hal-01128128

Submitted on 4 May 2016

PSYCHOACOUSTIC INVESTIGATION OF AUDITORY CUES INVOLVED IN HUMAN UNDERWATER SOUND LOCALIZATION.

Sophie Savel and Carolyn Drake

Laboratoire de Psychologie Expérimentale, CNRS UMR 8581, Université René Descartes, 71
av Edouard Vaillant, 92774 Boulogne Billancourt, France. savel@psycho.univ-paris5.fr

Theoretical arguments about underwater sound localization predict that auditory cues used in air are impaired in water. However, long term acclimatization could emerge due to exposure to the environment. We have compared localization abilities of expert and novice divers. The localization task was conducted for 8 azimuthal sound positions and 8 signals (0.4, 1, 2, 4, 6, 8, 10 kHz, white-noise). Results indicate that localization was better for experts: they made more correct localizations and their directional response distribution was less dispersed. These results demonstrate long term acclimatization in the processing of localization cues under water, particularly of interaural cues.

Introduction

Whereas auditory localization in *air* has been thoroughly investigated, the attempts to study this ability in water are limited. Meanwhile, the possibility for scuba divers to use acoustical signals for navigation systems would be of clear interest. Indeed, vision, the main sensory canal for spatial orientation in air, is severely impaired in water: 1) the width of the visual field is limited by the facemask, 2) deposits in suspension and the reduced spectrum modify the depth of focus, and 3) the lens effect produced by the glass of the facemask creates a distortion of the size and distance of objects. Thus, audition, the primary canal for communication and emergency signals, could acquire in water the status of a valuable canal for spatial orientation. The study of underwater auditory localization mechanisms involves, as a starting point, a rigorous knowledge of :1) the auditory cues that aid localization in the normal environment –air, and 2) the main physical properties of water that may distort human hearing mechanisms.

The ability to identify the position of a sound source in air is based on three acoustic cues: 1) interaural level differences between the signal reaching the two ears, 2) interaural time of arrival and phase differences, and 3) spectral cues induced by the pinna and head. The first two are *binaural* cues, because they involve the comparison of information from the left and right ears. The third localization cue is *monaural*, because it concerns the processing of the acoustical information from only one ear. Therefore, a left ear/right ear comparison involves primary interaural level differences (ILD), and time and phase differences (ITD). ILD occur because sound waves are diffracted by the head, so that the signal reaching the opposite ear is less intense. ITD both correspond to a difference in phases and to a difference in arrival time. Both ILD and ITD have a maximum value (0.63 ms, and 15 dB, respectively) for the most lateral positions (at the side of the ears). These values decrease linearly for sources located further back or forward from the ear axis, to a value of zero in the median axis (in front and in the rear, respectively). ITD dominate at low frequencies, related to a phase interaural ambiguity produced by short wavelengths. IID are dominant at high frequencies, because the head shadow effect operates only for wavelengths inferior or equal to the head diameter. A cross-over frequency has been established at 1.5 kHz [1]. Monaural localization is also possible: the external ear contains many resonance sites, so that some portions of the frequency spectrum of the signal are enhanced, while others are attenuated. Moreover, pinna produce incoming signal reflections, so that the time delay of arrival between the direct and reflected source strongly depends on the source direction [2]. However, specific resonance and reflection appear for frequencies above 3 kHz [3].

Human listeners are very precise in their ability to localize on the basis of these cues: the Minimum Audible Angle paradigm (MAA), that evaluates the minimum angular displacement that a listener can detect, is situated between 1° and 10°, depending on the source position and frequency [1]. However, studies measuring performance in an azimuth identification task indicate that even if listeners are easily able to identify the source angular separation, frequent front-back confusions are observed: when the source is presented in the frontal hemifield, listeners have a strong tendency to locate it in the rear hemifield, but at the same angular separation from the midline that the actual source azimuth, and vice-versa [4]. This front-back ambiguity is inherent to the equivalence of ID value for each source azimuth situated on a given "cone of confusion": ID values are symmetrical for azimuths θ = 0° and 180°; for θ = 15° and 165°, etc. Thus, listeners are unable to make a reliable front-back decision on the exclusive basis of binaural cues. This front-back discrimination is based on the processing of monaural pinna spectral cues, but only at high frequencies. The pinna influence on localization acuity has been demonstrated by selectively occluding different pinna cavities: the localization error rates appeared to increase with increasing pinna occlusion [5]. Moreover, the front-back confusion rate is constant and high at low frequencies, and decreases linearly for frequencies above 2 kHz [6]. The role of listeners' motion in localization has also been investigated. Head movements executed by the listener during the sound onset strongly reduce the front/back inversion pattern [7]. In summary, ITD and IID distinguish lateral position, spectral cues play a major role in front/back discrimination. ITD are dominant for low frequencies, IID and spectral cues for high frequencies. Head movements help for the front/back discrimination.

All studies on underwater sound localization start from the common assumption that the aerial cues described above are severely impaired in water [8, 9]. This assumption is based on physical differences between air and water, such as celerity and impedance. Water celerity (1435 m/s) is four times that of air (343 m/s). Thus, the value of ITD under water is reduced by at least four. In the same way, the impedance of water is much higher ($1.5 \bullet 10^6$ Pa • s/m) than that of air (428.5 Pa • s/m). Thus, because of the impedance mismatch that operates between water and head, the head is acoustically transparent and the ILD are reduced. In the same way, impedance relations implied in the pinna/water interface are unpropitious to wave reflection, so that spectral cues are lost. These theoretical arguments lead to pessimistic predictions concerning human sound localization. Moreover, the transmission of the sound energy to the inner ear is, in water, essentially assured by bone conduction of the skull and torso, rather than by eardrum and middle ear mechanisms. So, underwater hearing thresholds are higher than aerial ones [10].

Following these arguments, human listeners should be unable to localize sound sources under water. Meanwhile, preliminary studies have revealed some localization ability, even if considerably inferior to that in air. For instance, the Minimum Audible Angle is three times higher in water than in air, but improves with training [8]. Localization performance measured in terms of percentage of correct localization is largely above chance and also improves with training [11]. Thus, previous predictions were too pessimistic. However, underwater sound localization ability is far from being thoroughly investigated. All the studies cited measured performance with: 1) azimuths located in the frontal hemifield and 2) experienced divers. Moreover, performance was expressed in terms of correct localization rates. Most important findings described in aerial studies have been deduced from directional analyzes of the entire response distribution. This led to provide qualitative findings about the spatial map that the listener has constructed, and about the type of localization cues that have been processed. For instance, the front-back confusion pattern and the existence of cones of confusions have been inferred from the analysis of the direction of incorrect localization. Thus, in a previous study of expert divers localization performance in water, we have tested selective effects of the stimulus frequency and of azimuth on response distribution [12]. Results indicated an equivalent distribution for all azimuths producing the same theoretic ID size, and a strong front-back confusion pattern. This allowed us to postulate the persistence of binaural cues under water. Moreover, the best performance was observed with low frequency stimuli, suggesting that this binaural information is temporal. Whatever its nature, the magnitude of ID in water is a fourth of that in air. For instance, the maximum ITD is 170 μsec under water. This value produced at an extreme lateral position in water corresponds to that produced at an azimuth of 20° in air. This distortion should have enhanced localization errors on divers. Such a distortion in localization response has not been observed in our previous study with expert divers. Thus, we propose that the auditory system can acclimate in the processing of binaural information. The best manner to demonstrate any adaptation is the comparison of naïve and experienced listeners in the same auditory task. Whereas it is impossible to find naïve listeners in air, it is possible in water: while expert divers' ears and head have been frequently emerged, novice divers have not been exposed. Thus, a distortion of the spatial map should be observed on novice divers but not on experts. In order to test this hypothesis, novices and experts were compared in a azimuth identification task conducted with several azimuth and signal conditions.

A. Method

1. Subjects: 10 divers (ranging in age from 17 to 45 years) with normal hearing have been recruited and divided in two groups as a function of their number of dives: 5 novices with a mean of 18 dives, and 5 experts with a mean of 1070 dives.

2. Apparatus: A PVC pipe cage was emerged. Eight transducers were placed at every 45 degrees (at azimuths of 0°, 45°, 90°, 135°, 180°, 225°, 270° and 315°), equally spaced, at a listener/source distance of one meter. Transducers were calibrated to produce the same SPL. Signals were controlled by a computer located on a boat, and transmitted to an amplifier and a divider, before being sent to one of the 8 transducers.

3. Stimuli: 8 acoustic signals were used: 7 sine waves (0.4, 1, 2, 4, 6, 8, and 10 kHz), and a white-noise (0-10 kHz). Signals lasted 3 seconds, with a 25 msec rise/decay time. Sound intensity was 110 dB SPL.

4. Procedure: The experiment was conducted in sea water, in the Mediterranean. The diving cage was immerged to a depth of 10 meters. Divers, who wore their own personal diving suits but no hood, descended to the cage and sat on it. Experimental signals were sequentially presented in one of the 8 transducers. Each of the 8 signals was presented 3 times in each of the 8 transducers, giving 192 randomized trials. Responses were given on a 8-button response box, with each button representing the position of a transducer. A visual signal placed on the box preceded each stimulus, indicating that the divers should hold their breath (in order to avoid the noise of bubbles produced by the regulator). Two seconds later, a 3-second acoustic signal was produced. Divers could breathe after the offset of the sound, and had no time constraints about their answer. They had to indicate, by pressing one of the 8 buttons, which transducer was generating the sound. Responses from the box were recorded on the computer, and displayed on the monitor screen. Subjects received no feedback about their answers.

B. Results.

The mean correct localization rates obtained for each group, signal and azimuth are presented on Table. 1. Fig. 1 represents the same rates averaged over azimuth.

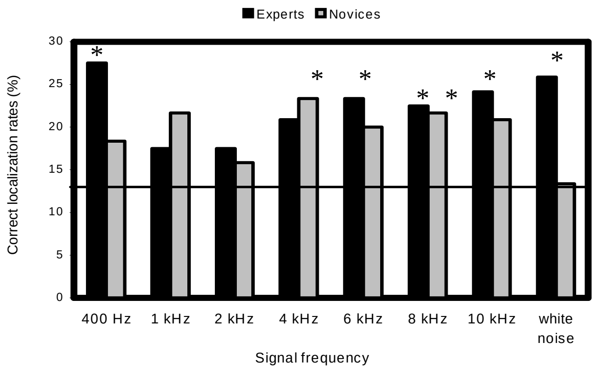

Fig. 1: Effect of divers' expertise on correct localization rates as a function of frequency, averaged over azimuth. The line indicates chance level (12.5%)

		Azimuth									
Frequency	*Group*	*0°*	*45°*	*90°*	*135°*	*180°*	*225°*	*270°*	*315°*	*Mean*	
400 Hz	*Experts*	33,3	0	20	20	73,3	13,3	40	20	*27,5*	***22.9***
	Novices	13,3	0	33,3	20	20	20	6,7	33,3	*18,3*	
1 kHz	*Experts*	26,7	20	20	13,3	13,3	0	33,3	13,3	*17,5*	***19.6***
	Novices	13,3	26,7	33,3	13,3	13,3	20	20	33,3	*21,7*	
2 kHz	*Experts*	26,7	13,3	20	33,3	20	0	20	6,7	*17,5*	***16.7***
	Novices	0	26,7	13,3	20	13,3	20	13,3	20	*15,8*	
4 kHz	*Experts*	13,3	6,7	6,7	33,3	33,3	26,7	26,7	20	*20,8*	***22.1***
	Novices	20	40	33,3	13,3	13,3	20	20	26,7	*23,3*	
6 kHz	*Experts*	33,3	26,7	26,7	26,7	6,7	13,3	40	13,3	*23,3*	***21.7***
	Novices	20	26,7	33,3	40	0	20	6,7	13,3	*20*	
8 kHz	*Experts*	40	6,7	33,3	13,3	26,7	13,3	33,3	13,3	*22,5*	***22.1***
	Novices	13,3	20	13,3	26,7	20	40	33,3	6,7	*21,7*	
10 kHz	*Experts*	6,7	6,7	33,3	53,3	20	6,7	26,7	40	*24,2*	***22.5***
	Novices	20	6,7	20	20	20	40	26,7	13,3	*20,8*	
white noise	*Experts*	13,3	13,3	20	33,3	26,7	46,7	40	13,3	*25,8*	***19.6***
	Novices	0	0	6,7	20	13,3	33,3	26,7	6,7	*13,3*	
	Mean	*18,3*	*15*	*22,9*	*25*	*20,8*	*20,8*	*25,3*	*18,3*	*20,9*	

Table. 1: Correct localization rates as a function of frequency, azimuth, and diver's expertise.

According to Fig. 1, the performance level strongly depends on frequency: for the 1 kHz and the 2 kHz stimuli, performance was poor for both groups, near chance level (12.5% in a 1/8 forced choice). Other stimuli were generally better localized. Regarding Table 1, the frequency effect is weighted by expertise effects. A repeated measures ANOVA carried out on correct estimates by azimuth (8), frequency (8), group (2) and repetition (3), revealed no effect of group. Meanwhile, by excluding 1 kHz and 2 kHz, there was a significant main effect of group [F(1,8)=6.3; p<.01]. The mean rate thus obtained is 24% for experts (significantly different from chance), and 19.5% for novices. Particularly striking are the results obtained with the 400 Hz sine and white-noise: performance level is very high for experts, with a large difference between the two groups.

In order to study a difference in the spatial map of the two groups, we have represented the entire response distribution obtained with the 400 Hz sine on Fig. 2. Data are represented as following: for each group, each of the eight stars corresponds to the response distribution obtained at a given azimuth (0° = star at the bottom; 90° = star on the left, etc.). The dotted line indicates the correct response rate; the full lines represent the other seven incorrect response rates. The longer the line, the more frequent the response in that direction.

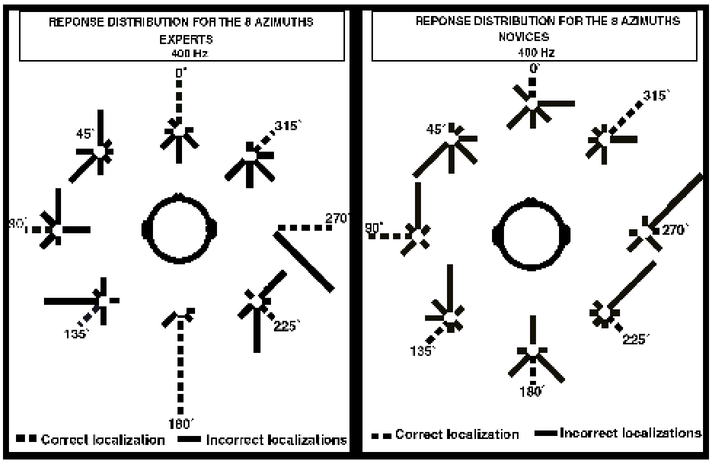

Fig. 2: Response distribution for the 400 Hz stimulus as a function of expertise and azimuth.

According to Fig.2, novice divers largely differ from experts in their localization response: experts responded more correctly, but their incorrect response are also less dispersed, and generally located at an azimuth adjacent to the actual one. On the contrary, novices made less correct localizations, and their errors are more dispersed. Whereas the mean correct rate is superior or equal to the average response rate at the two adjacent azimuths for experts, it is not the case for novices. Thus, experts seem to have made a better decision between two adjacent azimuths, while there was an ambiguity for novices.

C. Discussion.

The main frequency effect observed on performance suggests that listeners have processed a localization cue that is frequency dependent, otherwise the 8 correct rates would have been equivalent and not different from chance. The fact that a floor-effect was observed while localizing sources at medium frequencies from 1 kHz and 4 kHz reinforces the hypothesis of the use of binaural cues: in air, interaural phase and time differences can be processed with wavelengths longer than twice the head diameter -thus, for frequencies inferior to 1.5 kHz. Interaural level difference are effective for wavelengths shorter than the head diameter, that is for frequencies above 3 kHz. Thus, there is a medium frequency range where no ITD nor ILD are sufficient. Localization performance obtained in air at this frequency range is poor. Considering that ID exist in water, because of the quadruple water celerity (fivefold for sea water), wavelengths are greater. Thus, the ambiguous frequency range is displaced in water. However, results indicated a strong effect of divers expertise, both on the overall performance level and on response distribution. Experts are better able to localize, and their incorrect response are not equally dispersed over the 7 incorrect response possibilities. Novices made fewer correct localizations because they frequently confused the actual azimuth with an adjacent azimuth. Thus, for novices, there is an ambiguity between the auditory cue magnitude produced at an azimuth θ and that produced at an azimuth $\theta \pm 45°$. This suggests auditory acclimatization in experts. A long term adaptation in the processing of interaural differences has already been suggested with ontogenetic arguments: during growth, the diameter of the skull progressively increases. Thus, the interaural distance as well as the head shadowing are constantly modified, thus changing the magnitude of interaural information. So, listeners have to adjust their spatial map to these changes in ID values. Underwater localization may be concerned by such an acclimatization. Given the high level of difference between performance of the two groups obtained at low frequency (400 Hz), one can suggest that interaural time differences were used. However, it may concern the inter-cochlear delay more than the time-of-arrival at the ears, given that in water the sound signal directly stimulates the cochlea. Thus, the inter-cochlear delay under water (estimated with an inter-cochlear distance of 10 cm and a water celerity of 1500 m/s) is 65 μsec at a 90° azimuth, and 45 μsec at a 45° azimuth. This 20 μsec difference is discriminated by experienced divers, not by beginners.

These results highlight the interest of further studies. For instance, head rotations help localization in air by producing binaural scanning. The present experiment has demonstrated acclimatization in expert divers in the processing of binaural cues. This adaptation in processing small interaural differences would be more efficient if coupled with head movements. In other terms, we have investigated here *sensorial* adaptation, corresponding to an acquired ability of the auditory system to resolve small underwater ID values. Further studies will attempt to demonstrate *behavioral* adaptation by investigating how experienced divers have developed motor hearing strategies.

REFERENCE

[1] **Mills, A. W.**, On the minimum audible angle, *Journal of the Acoustical Society of America*, 30, pp. 237-246, 1958.

[2] **Batteau, D. W.**, The role of the pinna in human localization, In *Proc. R. Soc. Lond. B*, 168, pp. 158-180, 1967.

[3] **Shaw, E. A., and Teranishi, R.**, Sound pressure generated in an external ear replica and real human ears by a nearby sound source, *Journal of the Acoustical Society of America*, 44, pp. 240-249,1968.

[4] **Wenzel, E. M., Arruda, M., Kistler, D. J., and Wightman, F. L.**, Localization using nonindividualized head-related transfer functions, *Journal of the Acoustical Society of America*, 94, pp. 111-123, 1993.

[5] **Gardner, M. B., and Gardner, R. S.,** Problem of localization in the median plane: effect of pinna cavity occlusion, *Journal of the Acoustical Society of America*, 53, pp. 400-408, 1973.

[6] **Stevens, S. S., and Newman, E. B.**, The localization of actual sources of sound, *American Journal of Psychology*, 48, pp. 297-306, 1936.

[7] **Thurlow, W. R., and Runge, P. S.**, Effect of induced head movements on localization of direction of sounds, *Journal of the Acoustical Society of America*, 42, pp. 480-488, 1967.

[8] **Feinstein, S.**, Acuity of the human sound localization response underwater, *Journal of the Acoustical Society of America*, 53, pp. 393-399, 1973a.

[9] **Hollien, H.**, Underwater sound localization in humans, *Journal of the Acoustical Society of America*, 53, pp. 1288-1295, 1973.

[10] **Brandt, J. F., and Hollien, H.**, Underwater hearing thresholds in man, *Journal of the Acoustical Society of America*, 42, pp. 966-971, 1967.

[11] **Stouffer, J., Doherty, E. T., and Hollien, H.**, Effect of training on underwater sound-localization ability, *Journal of the Acoustical Society of America*, 57, pp. 1212-1213, 1975.

[12] **Bovet, P., Drake, C., Bernaschina; F. and Savel, S.,** Underwater sound localization: Role of interaural time and intensity differences. *Journal of the Acoustical Society of America*, 103(5), pp. 2844, 1998.