Investigation of auditory cues involved in human underwater sound localization

François Bernaschina, Sophie Savel, Pierre Bovet, Carolyn Drake

Investigation of auditory cues involved in human underwater sound localization

François Bernaschina*, Sophie Savel**, Pierre Bovet* and Carolyn Drake**
**Laboratoire de Psychologie Cognitive, Université de Genève, Genève, Suisse.*
***Laboratoire de Psychologie Expérimentale, CNRS UMR 8581, Université Paris V, Boulogne, France.*

Abstract : Theoretical arguments about underwater sound localization predict that auditory cues used in air are impaired in water, resulting in a theoretical inability to locate sounds under water. An azimuth identification task was conducted under water for 8 positions (0°, 45°, 90°, 135°, 180°, 225°, 270°, 315°) and 3 signals (400 Hz sine, 6 kHz sine, white-noise). Results demonstrate that localization was possible, at least for lateral positions, indicating that interaural cues are processed. The absence of spectral pinna cues may explain a large front/back confusion pattern.

INTRODUCTION

Whereas studies on human sound localization acuity are numerous, attempts to evaluate this capacity in water are rare.

In air, mechanisms underlying the ability to localize a sound source in the horizontal plane are now clearly identified. This sound localization ability depends on three acoustic cues: (1) interaural and phase differences between the signal reaching the two ears (ITD), (2) interaural intensity differences (IID), and (3) spectral cues induced by the head and pinna. The relative weight of these cues varies depending on the sound source position and its frequency components. Both ITD and IID have a maximum value (0.63 ms, and 15 dB, respectively) for extreme lateral positions (90° and 270° azimuth, at the side of the ears). These values decrease monotonicaly for sound positions further back or forward from this ear axis, to a value of zero in the median axis (0° and 180° azimuth, in front or behind the head). The dominance of one type of interaural difference depends on the frequency components of the sound to localize: ITD dominate at low frequencies, related to the temporal resolution capacity of the auditory system. IID are dominant at high frequencies, depending on the ratio between wavelength and head diameter. IID are due to a head shadow effect, resulting in an attenuation of the signal reaching the furthest ear. This head shadow effect operates only for wavelengths inferior or equal to the head diameter, that means for high frequency signals (Kistler & Wightman, 1992). A cross-over frequency has been established at 1.5 kHz, below which ITD are dominant, and above which IID are the most efficient (Mills, 1958).

In the median plane, both interaural values are equal to zero. In this case, the localization ability is based on spectral cues. These cues given by reflection properties of the pinna are efficient for frequencies above 2 kHz (Shaw & Teranishi, 1968; Gardner & Gardner, 1973). Spectral cues are involved in front/back discrimination. Under 2 kHz, frequent front/back confusions are observed (Stevens & Newman, 1936). Head movements executed by the listener during the sound onset strongly reduce this front/back inversion pattern (Thurlow & Runge, 1967).

In summary, ITD and IID distinguish lateral position wereas 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 optimize the front/back distinction.

All studies on underwater sound localization start from the common assumption that the aerial cues described above are severely reduced in water (Feinstein, 1973a; 1973b; Hollien, 1973; Wells & Ross, 1980). This assumption is based on physical differences between air and water, such as celerity and impedance. Water celerity (1435 m/s) is roughly quadruple that of air celerity (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 ($1.5 \cdot 10^6$ Pa • s/m) than air (428.5 Pa • s/m) (Rossi, 1986). Thus, the impedance mismatch that operates in air for head shadowing is altered in water. Thus, underwater IID are severely impaired, given that the head is acoustically transparent. 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. Also, head movements are restricted by the diving equipment. Moreover, sound transmission to the inner ear is in water essentially assured by bone conduction of the skull and torso, rather than by the eardrum and middle ear mechanisms. So, underwater hearing thresholds are higher than aerial ones (Brandt & Hollien, 1967).

Following these arguments, human listeners should be unable to localize sound sources under water. It is obvious interest to investigate this ability, given the applied aspect of relevant findings, for amateur and professional scuba divers in numerous fields (homing, communication, security...). Preliminary studies have revealed some localization ability, even if considerably inferior to that in air. For instance, the Minimum Audible Angle (minimum sound angle displacement that a listener can detect) is three times higher in water than in air, but improves with training to reach a level comparable to that of some marine mammals (Feinstein, 1973a; 1973b). Localization performance measured in terms of percentage of correct localization (35%) is largely above chance (estimated at 20% in an 1/5 azimuth identification task), and improves (to 50%) with training (Stouffer, Doherty & Hollien, 1975). Thus, previous predictions were too pessimistic.

The aim of the present study was to evaluate human localization performance by testing the entire horizontal plane and several signal frequencies, in order to determine how the various localization cues may be specifically implied and interact. Indeed, previous studies evaluated the main performance level averaged over sound position, and essentially tested azimuths located in the anterior part of the horizontal plane. By conducting here an azimuth identification task in the entire horizontal plane, and with several signal frequencies, we hope to distinguish the specific role of each localization cue (ITD, IID, and spectral cues). For instance, we predict that some use of interaural cues operates for the maximum size of ID values, that means for lateral positions. On the contrary, performance measured in the median plane, in the anterior and posterior part of it, should reveal that spectral cues are not involved. Subjects should make numerous front/back confusions.

METHOD

1. Subjects: All 14 participants (12 males and 2 females, ranging in age from 26 to 38 years) were experienced divers with an average of 600 dives, representing at least 75 hours spent underwater. They had all practiced diving for more than five years. They had normal-hearing between 125 Hz and 8 kHz.

2. Apparatus: An aluminum diving cage was constructed, with a saddle positioned in the center, in order to allow a precise sitting position for the subjects. A potentiometer was placed at the top of the hood in order to record head movements (max. ±121°). Eight aluminum arms were placed around the cage every 45 degrees (at azimuths of 0°, 45°, 90°, 135°, 180°, 225°, 270° and 315°, clockwise). Eight Motorola piezoelectric transducers were fixed at the top of each arm, equally spaced, in order to keep the listener/source distance constant at one meter. These 8 transducers responded to omnidirectional criteria and were calibrated to produce the same SPL intensity. Signals were controlled by a computer located on a boat, and transmitted to an amplifier (Monacor) and a divider, before being sent to one of the 8 transducers. All stimuli parameters (i.e., experimental signal, duration, attenuation, 1/8 transducer activation, sequential trials randomization) were controlled from the computer.

3. Stimuli: Three acoustic signals were used: 2 pure tones (400-Hz and 6-kHz sine waves), and white-noise (0-10 kHz). Signals lasted 2 seconds, with a 50 msec rise/decay time. Sound intensity level was 110 dB SPL.

4. Procedure: The experiment was conducted in Lake Leman, Switzerland. The diving cage was immerged to a depth of 10 meters, attached to a buoy by several cables in order to keep the system motionless. Divers, who wore their own personal diving suits with a 7 mm Neoprene™ hood, descended to the cage, sat on the saddle, and locked their legs by placing their fins beneath the cage. Each subject successively go through two experimental conditions: head movements were allowed in one condition and forbidden in the other. Experimental signals were sequentially presented in one of the 8 transducers. Each of the 3 signals was presented 3 times per condition in each of the 8 transducers, giving 144 randomized trials. Responses were given on a round 8-button response box, with each button representing the position of a transducer. A visual signal (placed both on the cage and 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). Three seconds later, a stimulus was produced. Divers could breathe after the offset of the sound, and had a 5-second delay to give their answer before the next visual signal. 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.

RESULTS

Contrary to our expectations, we haven't found an improvement of performance for the condition with head movements versus without head movements (Bovet *et al.*, 1998). This allowed us to regroup the results of both conditions for further analysis.

First, some localization ability was demonstrated. Performance measured in term of number of correct localisations is presented on Figure 1. Performance exceeded chance level (12.5%) for five of the eight azimuths, as tested by a t-test (<.05).

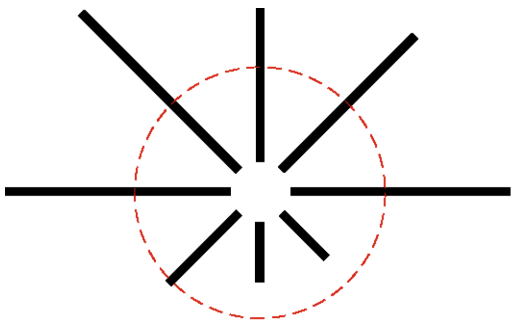

Figure 1: Total number of correct responses for the 8 azimuths. The dotted circle represents the chance level.

However, performance was not excellent (only about 30% correct in the best cases). A repeated measures ANOVA carried out on correct localization rates by azimuth (8), and signal frequency components (3) revealed that correct localization rates varied with azimuth [$F(7, 91) = 8.64$, $p < .001$]: correct localization rates were more frequent for sources located in front or laterally to the subjects' head (0°, 45°, 90°, 270° and 315° - mean = 26.7 % - double of chance level) than for those located behind it (135°, 180° and 225° - mean = 11.8% - not significantly different from chance).

Note that the mean for median positions (0° and 180° = 14%), is very much less than the mean for lateral positions (90° and 270° azimuth = 29.4%) as observed in air in the same identification task (Müller & Bovet, 1999).

To further examine the effect of azimuths, we investigated the way subjects responded, both when they identified the correct location, and when they were incorrect. Figure 2 presents three response distributions, one for each of the three signal frequency components. For each response distribution, data are represented at the eight azimuths (0° at the top, 180° at the bottom etc.). For each azimuth, the bold line indicates the responses in the correct direction, the grey lines represent the responses in the other seven (incorrect) directions, and the dotted arrows indicate the mean vector of each circular distribution.

An ANOVA was run on the absolute value of the angular difference between the directions of the stimulus and the response. The main results were:

First, white-noise and 400 Hz sine signals were located better than 6 kHz sine signal [$F(1, 13) = 11.29$, $p < 0.01$], with no difference between white-noise and 400 Hz sine signals [$F(1, 13) = 0.08$, *NS*].

Second, very few left/right confusions were observed (7% of incorrect estimates). This robust result indicates that whereas correct localization performance was poor (at best about 30% correct, see Figure 1), subjects were remarkably good at identifying whether a sound came from the left or right.

Third, anterior/posterior confusions were frequently observed when responding incorrectly. When the sound came from behind (135°, 180°, and 225°), 71% of incorrect responses were located in the forward quadrant (45°, 0°, and 315°). Inversely, when the sound came from in front (45°, 0°, and 315°), 45% of incorrect responses were located in the posterior quadrant (135°, 180°, and 225°). These confusions were asymmetrical with more posterior/anterior than anterior/posterior confusions (71% and 45%, respectively).

Moreover, a particular response pattern of back/front inversions appeared for the two azimuths located in the median axis (0° and 180°): when the sound came from the front (0°), incorrect responses were mainly located directly behind (180°), and when the sound came from behind (180°), incorrect responses were located exactly in front (0°), the latter being more prevalent (48% front/back confusions, 26.7% back/front confusions, respectively) . This systematic error could be related to the front/back confusion pattern classically observed in aerial experiments. Front/back inversions have been attributed in air to the absence of spectral cues.

DISCUSSION

These results first indicate that underwater localization performance level is higher than it could be expected on theoretical basis.

Second, there is a strong distinction between performance in lateral axis and median positions. Subjects are better able to localize a source as coming from a lateral position. Azimuths giving the maximum level of performance are those that give the maximum interaural difference value. This result supports the hypothesis that the maximum interaural difference value in air is sufficient but not necessary : in air, listeners

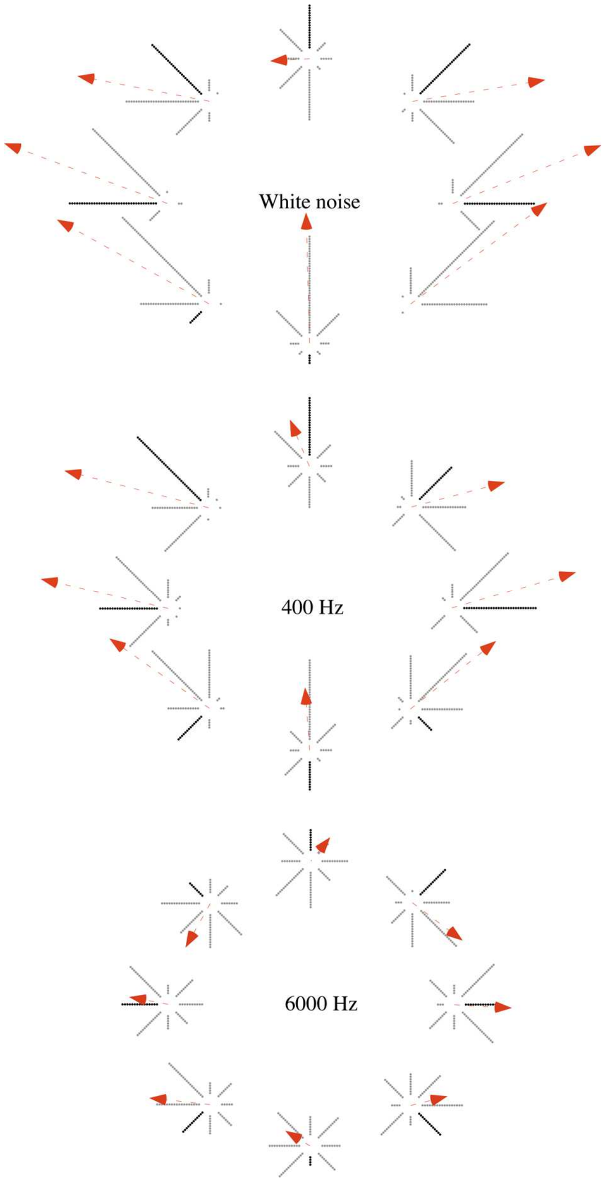

Figure 2 : Overall circular distribution of responses for the three acoustical signals. The bolded lines correspond to the correct responses, and the dotted arrows indicate the mean vector of each circular distribution.

may be able to localize (on the basis of interaural differences) with a value considerably smaller than the maximum stimulus size. Even if the maximum ITD size is 140 microseconds (for a head diameter of 0.20 m and a water celerity of 1435 m/s) under water, the inter-cochlear phase discrimination threshold in air is 6 microseconds (Tobias & Zerlin, 1959). Thus, the human auditory system should be able to process interaural differences under water, at least at maximum values, that is for extremely lateral azimuths. Nevertheless, the efficiency of ITD is limited by the signal frequency: there is an ambiguity as soon as the ITD approches the value of the signal's period. We saw that this frequency limit is around 1.5 kHz in air; and it can be estimated around 6 kHz in water. This could explain the poorer performances observed with our 6 kHz signal compared to the other signals.

For sources located in the median axis, performance was poor, and a strong front/back confusion pattern was observed. This confusion pattern, frequently examined in aerial localization, is in air related to the absence of interaural difference in the median plane, as well as the insufficiency of monaural reflection cues in low frequency ranges. In water we can consider that both IID and spectral cues disappear for all frequencies according to head/water impedence relation.

Moreover, Neoprene™ is known to attenuate sounds above 1 kHz (Norman, Phelps & Wightman, 1971), thus the worse performance for the 6 kHz might be related to a weaker signal. This points out the interest of further investigations of localization acuity, carried out without the hood. These further experiments without hood should test higher frequencies, in order to determine whether frequencies spectral cues can be processed under water, in which case the front/back inversion pattern should be reduced.

In the same way, these investigations conducted with a bare head with high frequencies should reveal the nature of the interaural cue used in the lateral plane. The cross-over frequency that separates the two interaural differences is 1.5 kHz in air. It is supposed to be displaced to 6 kHz in water (Hollien, 1973). The use of frequencies above 6 kHz should allow the evaluation of the role of IID under water.

REFERENCES

Bovet, P, Drake C., Bernaschina, F. & Savel, S. (1998) Underwater sound localization: role of interaural differences, *J. Acoust. Soc. Am.*, ***103** (5)*, 2844.

Brandt, J. F., and Hollien, H. (1967) "Underwater hearing thresholds in man", *J. Acoust. Soc. Am.*, ***42***, 966-971.

Feinstein, S. (1973a) "Acuity of the human sound localization response underwater", *J. Acoust. Soc. Am.*, ***53***, 393-399.

Feinstein, S. (1973b) "Minimum audible angle underwater: a replication under different acoustic and environmental conditions", *J. Acoust. Soc. Am,* ***54****,* 879-881.

Gardner, M. B., and Gardner, R. S. (1973) "Problem of localization in the median plane: effect of pinnae cavity occlusion", *J. Acoust. Soc. Am.,* ***53****,* 400-408.

Hollien, H. (1973) "Underwater sound localization in humans", *J. Acoust. Soc. Am,* ***53****,* 1288-1295.

Kistler, D. J., and Wightman, F. L. (1992) "A model of head-related transfer functions based on principal components analysis and minimum-phase reconstruction", *J. Acoust. Soc. Am.,* ***91****,* 1637-1647.

Mills, A. W. (1958) "On the minimum audible angle", *J. Acoust. Soc. Am.,* ***30****,* 237-246.

Müller, B. & Bovet, P. (1999) Role of pinnae and head movements on localizing pure tones, *Swiss J. of Psychology,* ***58.***Norman, D., Phelps, R., and Wightman, F. (1971) "Some observations on underwater hearing", *J. Acoust. Soc. Am.,* ***50****,* 544-548.

Rossi, M. (1986) *Traité d'électricité, Electroacoustique*, volume XXI, Lausanne, Presses Polytechniques Romandes.

Shaw, E. A., and Teranishi, R. (1968) "Sound pressure generated in an external ear replica and real human ears by a nearby sound source", *J. Acoust. Soc. Am.,* ***44****,* 240-249.

Stevens, S. S., and Newman, E. B. (1936) "The localization of actual sources of sound*", Am. J. Psychol.,* ***48****,* 297-306.

Stouffer, J., Doherty, E. T., and Hollien, H. (1975) "Effect of training on underwater sound-localization ability*", J. Acoust. Soc. Am.,* ***57****,* 1212-1213.

Thurlow, W. R., and Runge, P. S. (1967) "Effect of induced head movements on localization of direction of sounds", *J. Acoust. Soc. Am.,* ***42****,* 480-488.

Tobias, J. V., and Zerlin S. (1959) " Lateralization threshold as a function of stimulus duration", *J. Acoust. Soc. Am.,* ***31*****,** 1951-1954.

Wells, M. J., and Ross, H. E. (1980) "Distorsion and adaptation in underwater sound localization", *Aviation, Space and Environmental Medicine*, 767-774.