BIO Web of Conferences **1**, 00026 (2011)
DOI: 10.1051/bioconf/20110100026

Ears on the hand : reaching 3D audio targets

Vincent Forma[*‡], Thomas Hoellinger[*], Malika Auvray[‡], Agnès Roby-Brami[*] and Sylvain Hanneton[*]
(*) Laboratory of Neurophysics and Physiology, Paris, France
(‡)LIMSI, Laboratoire d'Informatique pour la Mécanique et les Sciences de l'Ingénieur, Orsay, France
E-mail: vince.forma@hotmail.fr or sylvain.hanneton@parisdescartes.fr

Abstract

We studied the ability of right-handed participants to reach 3D audio targets with their right hand. Our immersive audio environment was based on the OpenAL library and Fastrak magnetic sensors for motion capture. Participants listen the target through a "virtual" listener linked to a sensor fixed either on the head or on the hand. We compare three experimental conditions in which the virtual listener is on the head, on the left hand, and on the right hand (that reach the target). We show that (1) participants are able to learn the task but (2) with a low success rate and high durations, (3) the individual levels of performance are very variable, (4) the best performances are achieved when the listener is on the right hand. Consequently, we concluded that our participants were able to learn to locate 3D audio sources even if their ears are transposed on their hand, but we found of behavioral differences between the three experimental conditions.

1. Introduction

Many studies have shown the feasibility of converting visual information into sounds [1][2]. However, the sensori-motor parameters involved in the learning of a new coupling between perception and action remain unknown as well. We propose to address this question in the framework of the learning of a task that involves the interception of a static auditory target with the hand by people exposed to a particular 3D immersive auditory environment. Participants wore headphones and motion capture sensors on the head and on the right hand. They were seated and had to intercept a virtual static target emitting a "fly-like" sound within their workspace in two experimental conditions. The first mode ("head mode") can be considered as the more "natural" one since a head sensor is used to drive the position and orientation of the virtual listener in the 3D auditory display. In the second mode ("hand mode"), the position and orientation of the virtual listener was driven by a sensor fixed on the right hand. Consequently, in this condition, the sound produced by the fly is perceived through "ears" fixed on the hand of the participant. The aims of the study reported here were (1) to investigate participants' ability to localize a source within a 3D visual-to-auditory environment, (2) to investigate whether performance depends on the emplacement of the sensor (hand versus head), and (3) to study the relationship between the movements and the level of performance of participants.

2. Experiment #1

2.1 Materials and methods

We used an Electromagnetic device (Polhemus) connected with a 3D audio rendering system (OpenAl) that provides auditory feedback of movements. The sound perceived depends on the position and orientation of a virtual listener relative to the position of the source.

In the head mode, the listener was attached to the head and the perceived sound was sensitive to the position and orientation of the head with respect to the sound source. It corresponded to a relatively natural situation. In the hand mode, the listener was attached to the right hand and the perceived sound was sensitive to position and orientation of the hand. This mode corresponded to an unusual situation in which the participant had to adapt in order to achieve the task.

For each of the two modes, 10 participants (4 females and 6 males, age range: 25-50 years) were instructed to "catch the fly". They were blindfolded and two sensors were fixed respectively on the back of their right hand and on the closed headphones worn by the participants. The trial also stopped after 31s cut-off representing no-success. The "fly" was presented 3 times in 9 different positions for each of the two modes. Half of the participants began with the hand mode while the other half began with the head mode.

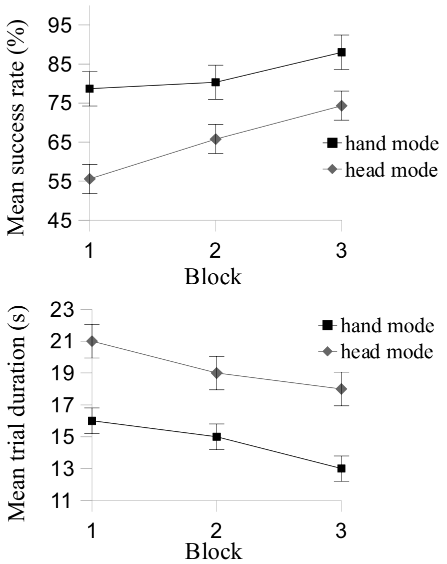

Figure 1: Evolution of the success rate and of the trial duration for each block and for the two mode. Error bars represent the confidence interval of the mean

2.2 Results

We calculated the success rate (percentage of catches before cut-off) and trial duration considered as a measure of the level of performance on the task. The kinematics of hand and head movements was characterized by the length of the trajectories, their mean velocities, and the cumulated rotations. The ratio of the length of the trajectory to the initial distance between the hand and the source (the "curvature") was also calculated as an index of efficiency of the movement.

We studied the influence of two repeated measures factors on the mean values of the variables cited above with ANOVAs. The first is the two-levels Mode factor (hand and head). The second is a three-levels Block factor. Each level of this factor corresponds to the number of the target presentation (i. e., first, second, and third time).

The Mode factor significantly influenced the success rate (percentage of catches before cut-off) ($F[1,9]=27.11$, $p<0.001$) and trial duration ($F[1,9]=16.67$, $p<0.003$). In the hand mode, it took the participants 14.55 s (s.e.m. = 0.569) to finish a trial with a mean success rate of 0.863 (s.e.m.=0.021) versus 19.26 s (s.e.m. = 0.661) with a mean success rate of 0.633 (s.e.m.=0.029) in the head mode (see figure 2). The index of efficiency was also significantly greater in the hand mode than in the head mode ($F[1,9]=9.48$, $p<0.014$). The mean jerk of the hand displacements was smaller in the hand mode (1450.4 cm/s3, s.e.m.=53.964) than in the head mode (2463.55 cm/s3, s.e.m.=141.591) suggesting smoother trajectories.

A significant influence of the Block factor was observed on the success rate ($F[2,18]=5.87$, $p<0.011$) and trial duration $F[2,18]=6.26$, $p<0.009$) demonstrating an adaptation to the task (see figure 1).

2.3 Discussion about the first experiment

Experiment 1 demonstrates that auditory feedback can be used in order to guide reaching movements. In addition, the analysis of the participants' success rate and trial duration revealed that participants' performance was immediately higher (from the first block) using the hand mode than using the head mode. We put forward the hypothesis that the larger number of degrees of freedom available to the hand as compared with the head allowed the participants to obtain more auditory information in order to locate the fly in the former case. Another possible assumption is a better coupling of auditory inputs with proprioceptive information associated with hand movement. In any case, it appears that perception increases movement quality and that inversely, movement increases perception.

The significant improvement which occurred both in hand and head mode with an increase in success rate and velocity demonstrates that healthy participants are able to adapt to a new audio-motor environment, as shown before [3] and/or to learn new efficient motor strategies.

3. Experiment #2

In Experiment #1, the participants had better performance in the hand mode than in the head mode. This result appears at first sight surprising given that perceiving an auditory source with the sensors located near the ears might appear as a more natural situation. However, this result might be due to the fact that the head condition involves more complexity. Indeed, in the hand condition, the virtual listener was located on the hand which was also used to catch the flight. In other words, the effector (i.e., the hand used to catch the flight) and the sensor (the listener) were spatially coincident. In order to investigate whether this spatial coincidence has an influence on the participants' results, we propose a second experiment (Experiment #2) that involves a hand condition in which the listener is on the left hand and where the right hand is used to catch the fly. The first aim of Experiment #2 was thus to investigate whether participants' level of performance remains better in the hand mode than in the head mode when the hand is not allowed to move toward the target and can only give to the participant a "point of view" of the target.

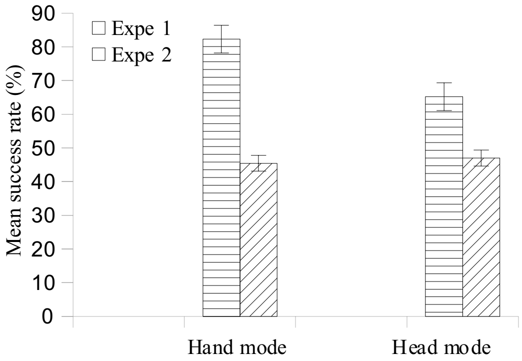

Figure 2: Mean success rate in hand and head mode for both experiments. Error bars represent the confidence interval of the mean

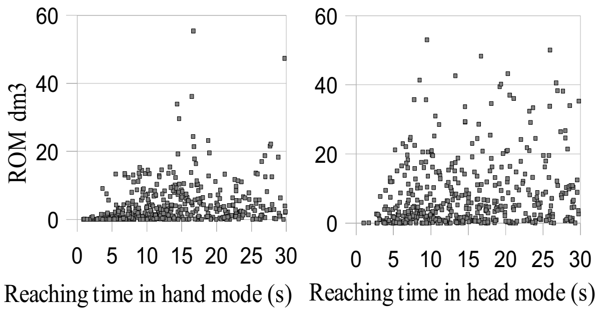

Figure 3: ROMs for reaching times

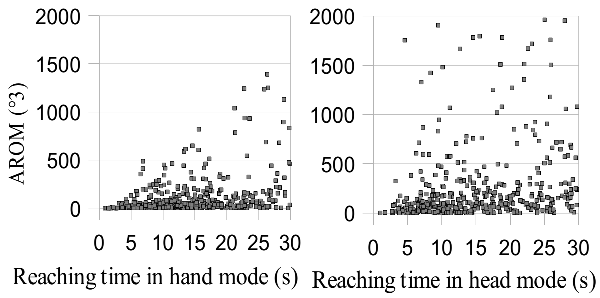

Figure 4: AROMs for reaching times

Target	0	1	2	3	4	5	6	7	8
hand mode	72	62	70	65	37	65	37	36	23
head mode	61	61	53	51	28	80	70	42	29

Table 1: Reaching (success) number for each target in both mode

3.1 Materials and methods

The same protocol was applied to 35 new participants (19 females and 16 males, age range: 20-50 years). Only the hand mode was modified. A third sensor was fixed on the back of left hand and participants were instructed to "catch the fly" with their right hand. In head mode the head supported the listener. Conversely in the hand mode, the listener was in the left hand. The elbow is fixed on the armrest of the armchair in order to limit the degrees of freedom of the left hand approximatively to those of the head.

3.2 Results

Contrarily to experiment 1, the level of performance is not significantly different between the two modes. Indeed it took the participant 15 seconds on average to reach a target with a mean success rate of 47 % in head mode and 14,2 seconds (45 %) in hand mode (see figures 2).

Whereas, while the mode factor doesn't affect distance traveled by virtual ears, it strongly affects their mean range of motion (ROM[1]) and angular range of motion (AROM). The mean ROM is 8,67 dm^3 in head mode that is significantly different from 5,15 dm^3 in hand mode ($p<0,0256$). The mean AROM in head mode is $516,73°^3$ and $219,75°^3$ in hand mode ($p < 0,0165$). Figures 3 and 4, that represents the ROM and AROM in both modes as a function of the reaching time, seems to illustrate a behavioral difference between the modes.

In the hand mode, large ROM or AROM were found only when the exploration times go higher than about 15 seconds. In order to explore the difference in the repartition of the AROM in the two modes, we distributed the AROM values in ten intervals. We obtained two significantly different distribution according to the mode (χ^2= 35.4405, df=9, $p< 0.0001$).

Another result seems interesting : in each mode some targets are more successfully caught than others depending their position in space. The reaching number for each target in each mode is sum up in table 1.

The number of target caught for each target differs significantly from a uniform distribution in both modes (head mode χ^2=46.83, dof=8, $p <0.0001$, hand mode χ^2 = 53.6516, dof=8, $p<8.065e-09$). Additionally the

1 In order to have an estimation of "how much" the head or the hand moves during a trial, we computed to variables named range of motion (ROM) and angular range of motion (AROM). The range of motion ROM is the volume (in dm3) obtained by the product of the range of motion of the sensor in the different axis directions (x,y,z). The angular range of motion (AROM) is the same but for the three Euler orientation angle of the sensor. It is the product of the range of motion for the elevation, the range of motion for the azimuth and the range of motion for the roll.

targets more often touched are not the same in each mode (χ^2= 20.4715, dof=8, p <0.009).

The Block factor influenced significantly the number of caught target, revealing an adaptation to the task (χ^2=6.0121, dof=2, p<0.05 for head mode and χ^2=8.2887, dof=2, p<0.016 for hand mode). The learning is significant between the block 1 and 2 for hand mode (p = 0.00677) and between the block 1 and 3 for head mode (p = 0.00703). (see figure 5).

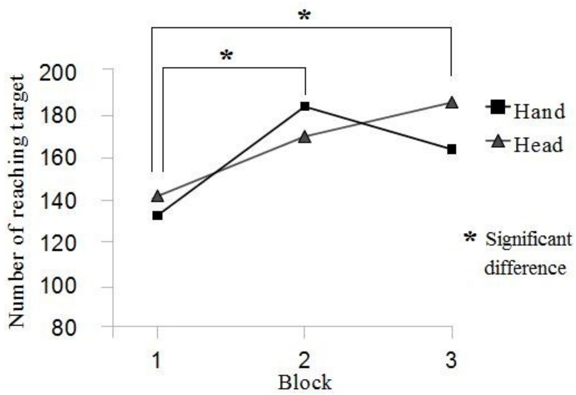

Figure 5: Number of reached target in each block for both modes

3.3 Discussion about the experiment #2

We did not find any significant difference between the two modes for the success rate and the trial duration. The experiment #2 seems to validate the hypothesis we made that the higher level of performance obtained in the hand mode in experiment #1 was due to the fact that the effector and the sensor were spatially coincident. In other words, in experiment #2, both perceptive tasks were equivalent whereas in the first experiment we compared the head mode to a kind of "guided approach" reaching task. In experiment #2 we show that participants were able to adapt their motor behavior to a new way of listening that involves a new sensory-motor coordination between the two hands. Surprisingly, the head mode that was supposed to be very close to a natural condition (for the human species, the ears are on the head) is not significantly better that this new listening mode, and improvements are slower. However, even if the level of performance seems to be similar, participants move their virtual ears with smaller amplitudes in hand mode.

If we consider the spatial aspects of the performance, in head mode the highest targets are more often reached than in hand mode (χ^2=16.69, dof=3, p<0.008). But we did not found any significant difference due to the mode factor concerning the rate of success for right versus left targets or near versus far targets.

4. General Discussion

We show that participants are able to learn the proposed tasks but with a low success rate and high durations. The low level of performance especially in the head mode can be explained by the fact that we used only the basic functions of the OpenAL library that do not involve for instance individual HRTF that would suit the head anatomy of the participants. However, our aim was to investigate the adaptation to a new sensory-motor coordination and not to produce a realistic 3D environment.

We found tremendous differences between participants : for instance in the hand mode for experiment #2 the best obtained a 92,6% success rate against a 7,4% success rate for the worst. The best performances are obtained when the listener is on the right hand i.e. when the listener and the effector are spatially coincident. In this condition, the task is rather a "guided approach" task than a pure reaching task. This result is very interesting : in a degraded perceptive situation (real or virtual), it could be more efficient to use spatially coincident sensors and effectors, with an immediate improvement of the level of performance.

The successful audio-motor coupling observed in these experiments suggests that this type of paradigm could be used to elicit movements by 3D auditory feedback. For instance, augmented auditory feedback in a game-like situation may be a useful tool for rehabilitation of sensory-motor functions [4].

References

[1] Auvray M., Hanneton S., et O'Regan JK. 2007. « Learning to perceive with a visuo – auditory substitution system: Localisation and object recognition with 'The vOICe' ». *Perception* 36 (3): 416-430.

[2] Bach-y-Rita, P. 2003. « Sensory substitution and the human–machine interface ». *Trends in Cognitive Sciences* 7 (12): 541-546.

[3] Säfström, D., and Benoni B. Edin. 2006. « Acquiring and adapting a novel audiomotor map in human grasping ». *Experimental Brain Research* 173 (3): 487-497.

[4] Robertson JV, Hoellinger T, Lindberg P, Bensmail D, Hanneton S, Roby-Brami A. Effect of auditory feedback differs according to side of hemiparesis: a comparative pilot study. J Neuroeng Rehabil. 2009 Dec 17;6:45.