AUDITORY VELOCITY INFORMATION IN A BALANCING TASK

Matthias Rath

Berlin University of Technology, Deutsche Telekom Laboratories
Ernst-Reuter-Platz 7, 10587 Berlin, Germany
matthias.rath@telekom.de

ABSTRACT

Within the general context of auditory perception of ecological information a previously rather less studied aspect is the one of the convection of continuous dynamic physical attributes. The study focuses on velocity information in a scenario of interactive control, the one of balancing a virtual ball on a tiltable track. In a target reaching experiment control movements and performance times are measured and recorded under different conditions of auditory feedback in addition to a wide–screen graphical display. The presence and relevance of auditory perception of velocity information can be inferred from analysis of experimental results and conclusions can be drawn concerning the design of auditory feedback of ecological or rather abstract nature.

[Keywords: Auditory feedback, Gestural control]

1. INTRODUCTION

1.1. Background

In the tradition of psychoacoustics, focus has for long been on the perception of abstract properties of sound such as pitch/frequency or loudness/level. The last decades however have seen an increasing interest in auditory perception as a means to gain ecological information from sound, i.e. information about physical objects and events in a listener's surroundings, through sounds they emit. As examples, studies have examined the auditory perception of physical attributes such as size [1] or the classification of auditory events as bouncing or breaking [2]. As well in the recent decades, the field of auditory display has strongly emerged and also here the question of the usage of abstract sounds with attributes related rather to traditional musical terms, such as in *earcons* [3], or of sound of rather ecological orientation, such as in *auditory icons* [4], is of central interest. While abstract sonic attributes leave more freedom in the design of auditory displays and may also facilitate the introduction of "grammatical" structures in the acoustic communication [5] [1], one may hope to reduce the need of explanation, training and adaptation on the side of the user by exploiting mechanisms of ecological auditory perception that are already internalised. This question of abstract vs. ecological auditory attributes is one central notion behind the work presented in the following.

Most psychoacoustic examinations of information convection through sound deal with discrete, often short, sonic signals or however with information conveyed in discrete events and classes. Just similarly, and this parallel is probably not by chance, in most human–computer interfaces sound is so far employed only (if at all ...) in the form of short discrete signals, typically of warning or notification. Continuous information flows through as well continuous sonic feedback, in particular in connection with gestural interaction, have in contrast hardly been subject of dedicated studies or employed in applications. This lack of attention stands in contrast to our everyday physical surroundings with which we generally interact in seamless continuous ways: we rather touch or move objects than toggle discrete switches or states and generally simultaneously receive continuous sensory feedback from the world. As a familiar example one may think of driving a car where the noises of motor, brakes or air streams continuously inform about processes to control and influence the driver's behaviour. The *Sounding Object (SOb)* European research project [6] for which the author has been working has been occupied with the development of (and psychoacoustic research as a basis for) *sound models*, sound generation algorithms that are able to convey in a clear and intuitive way, to non-expert listeners, ecological information contained in everyday sound-emitting scenarios. These sound generation algorithms can react dynamically to realtime parameters such as user input and reflect evolving states of a system to interact with and can therefore be seen as a dynamic, reactive extension of Bill Gaver's concept of *auditory icons*.

Related to the mentioned aspect of continuous sonic feedback is a principal problem of assessing perceptual mechanisms which may possibly happen unconsciously. Most conventional techniques of psychophysical evaluation, such as questionnaires, rating, scaling or labelling tasks, rely on conscious reflection and answers of the test subject. Sensory–motor mechanisms may however happen without a subject's awareness and possibly even resist any assessment based on conscious reactions: such experiments provoke, and rely on, a chain of self-observation and cognitive reflection on the experimental task (or request) as well as the own response, which may generally introduce side-effects that put into question the value of the experimental results in the light of the actual point of interest. The study described in the following represents an alternative approach that allows to by-pass the described problem. Here, movement responses of test subjects are measured and analysed, in a sensory–motor task that does not request any conscious judgement of sonic attributes. From correlations between characteristics of measured control movements and given conditions of sensory feedback, facts of perception of information through auditory feedback may be inferred, also if these are located beyond subjects' conscious awareness.

1.2. Metaphor and interface

The *Ballancer* is a tangible audio–visual interface designed to examine questions of the context described above. It is based on the metaphor of balancing a ball rolling along a tilt-able track. The device is handled by the user as if balancing a small marble on top

[1] These are probably the reasons why abstract sounds dominate the field of *sonification*.

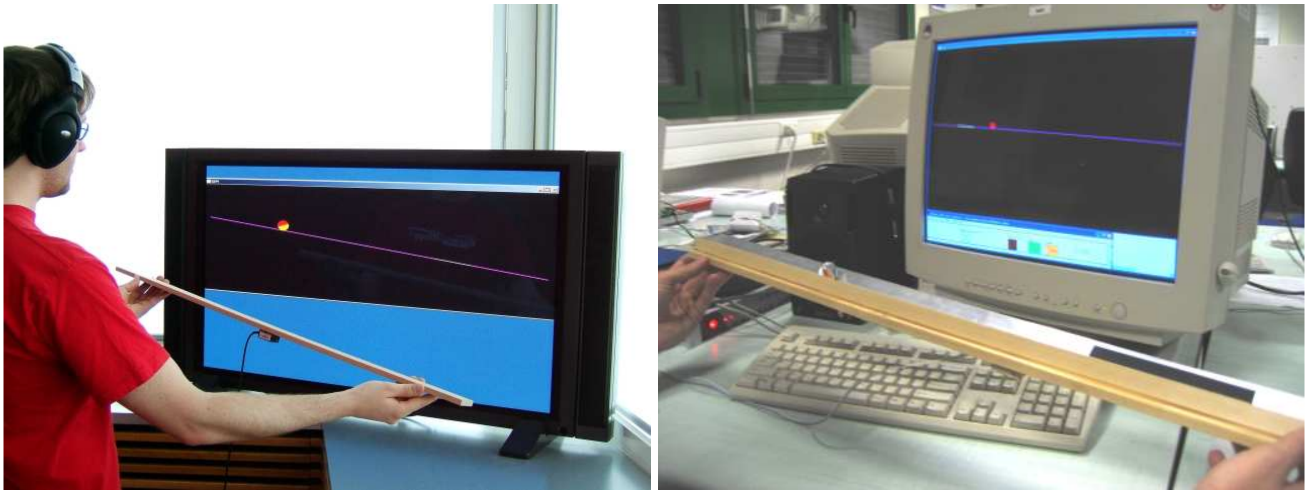

Figure 1: *The* Ballancer *in the configuration with a wide-screen display spanning the whole size of the 1–*m *physical control stick (left) and in the setup of the previous study with a 19" monitor (right). The real glass marble on the track in the right–hand photo only serves demonstration purposes.*

of a 1-m long wooden track whereby feedback about the movement of the virtual ball can be given through different types of visual and/or auditory feedback. Figure 1.2 shows two photos. In its original main configuration the *Ballancer* makes use of a *sound model* of a solid object rolling along a plane [7] which forms one example for the development of *sound models* that this author has been occupied with in the course of the *Sounding Object* European project [6]. The experimental setup used in the experiment described in this article (in the following sections) includes a second, more abstract approach of sonic feedback. For the visual feedback the first realisation of the *Ballancer* used a standard 19" computer monitor (see right photo of figure 1.2) and the first experimental study conducted at the device (shortly summarised in the next subsection) used graphical displays of different sizes, from the largest filling the whole 19" screen to the smallest spanning $1/12$ of this size.

Core of the gestural input is a sensor connected to the wooden control track that allows the measurement of the track's angle with the horizontal. The experiment presented in the following has been accomplished on a setup based on an *Inertia Cube* [8] orientation sensor.

The movement of the virtual ball in dependence of the changing inclination of the control track is realised through an implementation of the discrete–time version of simplified physical equations describing the underlying scenario. Theoretical and practical details behind the *Ballancer* interface have been described in a dedicated article [9]. Despite simplification of physical details the reactive behaviour of the virtual ball, in connection with the sound of the *rolling model* [7], was seen to be perceived as convincing and the metaphor was found to be understood spontaneously by test subjects without any prior explanation (as explained in the next subsection and described in detail in [9]).

1.3. Previous work and results

An initial study with 10 test subjects had been conducted at the *Ballancer* interface whose results are basis and motivation for the work presented here and shall therefore be shortly summarised. Here only main points are listed, details can be found in dedicated publications [10][9]. 10 subjects participated in this experiment containing as its main part a task of moving the virtual ball inside a fixed target area under the different feedback conditions of graphical displays of different sizes (the largest filling the 19" screen) and with and without sound feedback from the *rolling model*.

- The sound of the *rolling model* by itself (without connected graphical or other sensory information) was seen to provoke in most subjects a spontaneous connotation of a scenario of rolling.
- When handling the interface blindfolded, receiving feedback only through sound from the *rolling model* [7] (apart from proprioceptive feedback of their own arm movements) test subjects generally understood the underlying metaphor without any explanation. While clearly having an artificial character the sound of the *rolling model* was nevertheless seen to be less ambiguous in representing a rolling object than the sound of a real glass marble rolling on the tilted (blindfolded…) track.
- The 10 subjects were found to conclude a task of stopping the virtual ball inside a target area in average significantly faster when feedback from the *sound model* of rolling was present than with purely visual feedback. This effect of improved performance with sound was stronger for graphical displays of smaller size but present also for the full size 19" display.
- The noted faster task performance with additional sonic feedback could be connected to significant differences in indices of qualities of control movement such as the time after which the controlled virtual ball reaches its maximal velocity. In this sense the performance improvement with sonic feedback must be assumed to be (at least partly) due to optimised behaviour of acceleration and stopping with sound.

2. SCOPES AND SETTING OF THE PRESENT STUDY

The experiment described in the following aims at questions inspired by the results of the study described in the previous subsection (1.3). It forms an extension of this study as well as a concentration of focus.

As one fundamental difference to the study cited above (subsection 1.3) the following experiment is not driven by interest in the relevance of sound for small display sizes or in sound as an alternative for weak display conditions in general. In this respect, interest is here wider, on the question if sound can contribute to the performance of a task of continuous sensory–motor human–machine interaction by conveying information that may *in principle* not or not as effectively be conveyed to the user through visual feedback alone. While this notion is discussed further below, it is already noted that one of its consequences is the use of a wide–screen display spanning the full range of the physical control stick of the *Ballancer* in the setup used in the following (as seen in figure 1.2). All measurements in the following experiment are therefore accomplished under visual conditions that are optimal in the sense of the visual feedback spanning the whole range of the subject's arms during task performance, as would be the case in the "real" physical scenario. Following the same direction of reasoning the graphical display differs from the one of the previous study in that the ball is here not monochrome but with two differently coloured halves (left photo of figure 1.2) so that the turning movement itself may serve as an additional visual cue of ball movement.

2.1. The question of velocity information from sound

For the field of auditory display it is natural to ask if there exist specific classes of information that may be displayed particularly well or "in a natural way" through sound, or that vice versa show a predestination to be conveyed through the visual channel. For example does visual perception show a generally higher resolution in terms of the angular direction of a source of stimulus so that positional information is in most cases probably displayed and perceived more easily through vision than sound. With respect to temporal resolution the situation however appears rather opposite since attributes of acoustic stimuli residing on a level of temporal resolution much below the scale of visual perceptibility may generally still lead to differentiations in auditory perception: graphical frame rates around $30Hz$ are generally sufficient for the display of seamless visual motion while in the acoustic domain sampling rates above $44100Hz$ are generally considered necessary for transparency of a sampling process. As a consequence, the auditory channel may from a viewpoint of ecological perception be conjectured to show strengths in the convection of dynamical physical attributes such as forces and velocities. This idea indeed seems reasonable also from informal every day experience, as e.g. in the already mentioned (in subsection 1.1) situation of driving a car, even if we generally seem to be much less aware of this aspect of auditory perception as compared to the noted relevance of visual perception for spatial orientation.

In the initial configuration of the *Ballancer* the sonic feedback roughly reflected (and therefore potentially informed about) the position of the virtual ball through simple amplitude stereo panning and in the fact that the target area was marked by a different surface structure and as a consequence different acoustic behaviour. In the connected study (subsection 1.3) it was however shown that the found performance improvement through sonic feedback was accompanied by optimisations of control movements already in the phase of the task before reaching the target area. Since positional information in amplitude panning is very coarse it must be assumed that the noted performance effects must be at least partly attributed to velocity information conveyed by the sonic feedback. The present work focuses on this latter aspect and to this end employs conditions of sonic feedback that *exclude any possibility of positional information in the sound*. The momentary velocity of the virtual ball is here used as the only input parameter. In particular, the sonic feedback does not depend on the controlled virtual ball's location in- or outside the target area.

2.2. Ecological vs. abstract sonic feedback — types of auditory feedback

The *sound model* of rolling [7] used in the initial configuration of the *Ballancer* is based on a simplified physical model of the inner resonance behaviour as well as the force acting between the rolling object and the plane to roll on. The involved surface profiles that are "scanned" during the rolling movement form here the initial source of vibration, that is processed as part of the *sound model* following considerations on the physical and geometrical principles of the scenario. (I refer to [7] for details.) The development of the model followed the notion of *cartoonification* [4][9], i.e. it aims at informativeness or "expressiveness" and clearness and simplification in its sonic appearance rather than realism. Success in these scopes was confirmed by the study reported in subsection 1.3 were the sound feedback from the model was, while being clearly recognisable as synthesised, found to be intuitively understood by subjects in the sense of supporting well the convection of the intended metaphor and handling as well as leading to improved performance at the interface.

After these observations the notion of further "abstracting" the sonic feedback and possible implications on the convection of velocity information is central to the experiment reported in the following. A second, very simple and rather abstract sound model has been derived, by widely ignoring any idea of realism or even immediate similarity with the mechanical sound of a rolling object. This sonic feedback however still aims at expressing in a possibly intuitive way what is considered the main parameter of interest [2] for auditory display in our context, the one of velocity of a controlled movement. To this end the processing that accounts for the physical interaction in rolling is strongly reduced and the model is "stripped down" to scanning a chosen surface profile – at "audio rate" [3]. This strategy may be compared to replacing the object *rolling* along the surface with an ideal needle as of a record player that follows what would be the "essential" trajectory of the movement (of the centre of the virtual ball). Finally, a highly artificial and rather unrealistic surface profile is chosen, one of the shape of a lowpass-filtered sawtooth signal, to optimise the low-level psychoacoustic properties of the resulting signal: The Fourier spectrum of the sawtooth spreads over a very wide range of the frequency sensitivity of human hearing, up to its upper limit even for fundamental frequencies at the lower end of the hearing range. The sawtooth is periodic and thus stimulates a clear and strong sensation of pitch, e.g. in contrast to filtered noise. The signal is slightly lowpass filtered in order to minimally smoothen the otherwise extremely harsh aesthetic appearance of the sawtooth. Summing up the consequences of the described derivation of the record–needle model, the fundamental frequency of the used saw-

[2]The *sound model* of rolling in contrast to our specific focus here offers a potential to acoustically reflect a wider range of ecological attributes of the underlying scenario, such as the size, shape and weight of the rolling object, surface structure or material.

[3]i.e. with a temporal precision high enough to represent a continuous process to the human auditory system, exactly: at "cd audio rate" 44,100 Hz. . .

tooth signal follows proportionally the ball velocity to be sonified. This condition of sonic feedback shall in the reminder for simplicity be denoted as "*abstract sound* (*as*)".

2.3. Experimental design

In order to examine the effects of the more abstract, very simple sonic feedback on users' handling and perception of the Ballancer device, as compared to the original sound model of rolling, a second experiment analogous to the one described above was conducted, containing both conditions of sonic feedback (again along with a condition of purely visual feedback). Thereby visual feedback is as described above (at the beginning of section 2), identical for the whole course of the experiment. 6 subject participated in this pilot experiment, all of them students at Berlin University of Technology, aged between 21 and 27, four male and two female. Despite this rather small number of participants (that is augmented in current experiments continuing the pilot) some statistically significant and interesting results were found, the most important of which are shortly reported below. Again (as in the initial experiment of subsection 1.3) participants were asked to move the virtual ball inside a graphically marked target area (compare figure 1.2) as fast as possible, this time under the three different conditions of *a)* sound feedback from the rolling model, "*rolling sound (rs)*", *b)* feedback from the more abstract record–needle model, "*abstract sound (as)*" and *c)* without sonic feedback, "*no sound (no)*". The task was counted as completed when the virtual ball completely stayed at rest inside the target area for at least $500ms$ (stopping condition, compare also figure 3). For this condition to be achievable by the controlling subject the computation of movement of the virtual ball includes a simple model of "stick" and "roll" friction. The different conditions appeared in *sets* of 20 *games* each. The order of the sets/conditions was counterbalanced across subjects and the whole series of all conditions was repeated once for each subject so that the whole test consisted of 6 sets, e.g. for subject 1 of the form: "*rs, as, no, rs, as, no*", subject 2: "*rs, no, as, rs, no, as*"... Due to the repetition of the whole series, each condition appeared twice for each subject, as one set in a less trained ("untrained") state and again in "trained" circumstances in the second half of each test. Table 1 gives a quick overview of the distribution of feedback conditions for the 6 subjects in the design of the experiment. By counterbalancing the order of conditions we

sub-	Sound conditions of sets					
	"untrained"			"trained"		
ject	1	2	3	4	5	6
1	*rs*	*as*	*no*	*rs*	*as*	*no*
2	*rs*	*no*	*as*	*rs*	*no*	*as*
3	*as*	*rs*	*no*	*as*	*rs*	*no*
4	*as*	*no*	*rs*	*as*	*no*	*rs*
5	*no*	*rs*	*as*	*no*	*rs*	*as*
6	*no*	*as*	*rs*	*no*	*as*	*rs*

Table 1: *Counterbalanced order of feedback conditions (*"rolling"*, *"abstract"* and *no sound*) in the 6 sets of all 6 subjects.*

can hope that training effects during performance of the test "even out" in the comparison of measurements at different conditions of feedback averaged over all subjects.

3. RESULTS

3.1. Task performance under different conditions of feedback and training

As in the initial experiment with the Ballancer in the $19''$-screen configuration (subsection 1.3), the 6 subjects in average performed the task faster with sonic feedback than without. Table 2 shows the "*task times*", i.e. times subjects needed to complete the task in average — over all 6 subjects and 20 games — under the different conditions of feedback, in the first ("untrained") and second ("trained") series of sets. One interesting new observation with

Average *task times* (*s*), training	*rs*	*as*	*no*
"untrained"	7.62	8.38	9.28
"trained"	7.14	6.61	7.36
$\delta(\%)$	-6.2	-21.1	-20.6
p	0.261	0.000	0.001

Table 2: *Times (in* s*) subjects needed in average to complete the task under the different conditions of feedback in the "untrained" (line 1) "trained" (line 2) series. The third line shows the relative difference (of the value in line 2 with respect to the one in line 1 of equal feedback condition, in %), line 4 holds the p-values resulting from t-test comparison of these according underlying sets of values.*

respect to the initial Ballancer experiment (subsection 1.3) is the fact that the average training effect, i.e. the improvement of performance in doing the task over time between the "untrained" and "trained" sets — depicted in line 3 of table 2 — is comparatively small (approx. 6.2%) with the *rolling sound* and does not reach a statistical significance of 5%. The training effect is however much higher ($> 20\%$) and about equally strong under the conditions of *abstract sound* and without sound feedback. [4] In accordance with this phenomenon, average performance in the trained series gets better with the *abstract sound* than with the *rolling sound*, while it is best with *rolling sound* in the untrained series. Figure 2 where the mean values of table 2 are depicted graphically allows an easy overview over the "global" temporal behaviour of *task times* under the different feedback conditions.

Table 3 shows a comparison of the mean performances under different conditions in the form of relative difference and according statistical significances, i.e. p-values resulting from t-test comparison of the two respective sets of measurements. It can be seen that in the untrained series task performance is significantly faster with *rolling sound* than without sonic feedback and still faster with *abstract sound* than without, but the latter improvement does not reach statistic significance. In the trained series *rolling sound* and *abstract sound* somewhat "switch roles": still, with both types of sonic feedback average performance is better than without sound but now the *abstract sound* leads to the best (significantly...) average *task time* (please compare also figure 2).

3.2. Indices of quality of control movement

As discussed in detail in the previous subsection the measurements in the experiment presented here reveal that again (i.e. as in the

[4] In table 2 as in all following, p-values below or close to a statistical significance of 5% are highlighted green.

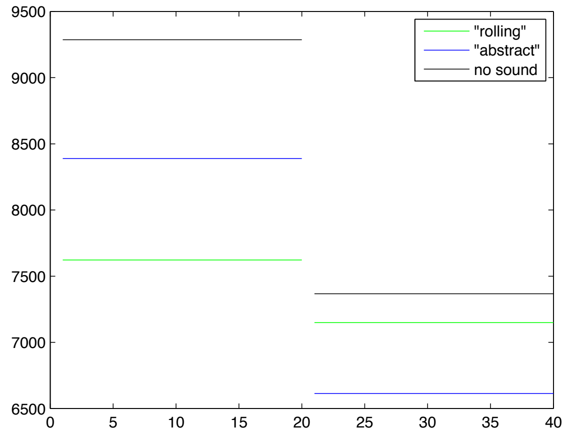

Figure 2: *Average* task times *(in* ms*) under the different conditions of sonic feedback, over the "untrained" (1-20) and "trained" (21-40) games and across all subjects. Note that the range of the y-axis does not start at zero for better visibility of the values.*

Relative differences of *task times*, statistic significance	"untrained"		"trained"	
	as	*no*	*as*	*no*
rs				
$\delta(\%)$	-9.1	-17.9	8.1	-3
p	0.1345	0.0052	0.1414	0.5977
as				
$\delta(\%)$		-9.7		-10.2
p		0.1825		0.0206

Table 3: *Differences in average* task times *(in %) under the different conditions in the untrained and trained set. Below each difference value, the according statistical significance, p, is given.*

previous *Ballancer* study, compare subsection 1.3 and [9]) subjects in average conclude the target reaching task faster with sonic feedback than without. This effect must here clearly be attributed to velocity information in the sonic feedback since the position of the controlled ball is not reflected in the sonic feedback in any way. It is natural to ask if or how this effect of in some way optimised control movements can be further qualified. A worthwhile goal would of course be to develop mathematical models of the control behaviour of the human operator that allow to quantitatively (or at least qualitatively) predict such effects of different feedback conditions in tasks of motor control and perception. Such models have a strong value in allowing to derive specifications for the design of human–machine interfaces and their development has since long been a topic in the context of control theory and robotic systems (see e.g. [11][12]). Yet, human operator [5] models that explicitly include a consideration of conditions of sensory feedback other than visual still form an area with open space for future work. Also in this publication no such human operator model that integrates the movement responses measured in experiments at the *Ballancer* is provided. The derivation and discussion of indices of measured movement trajectories that may represent characteristics of control behaviour may be seen as one step in the direction just hinted at. Characteristic points and phases in the movements of control stick and controlled virtual ball and connected indices are identified.

Figure 3 shows an example of temporal development of angle of control track (red) and position of controlled virtual ball (blue) in one game of the experiment. The left and right boundary of the target area are marked in green on the y–axis (i.e. by the horizontal green lines at $0.7m$ and $0.8m$). It is remarked that the angle is here represented by its sine value (stretched by a factor 2 for reasons of better visibility) which is in this range, with an absolute value below about 0.1, approximately proportional to the angle in radian itself. (A radian of 0.1 corresponds to approx. $5.7°$.) Several characteristics of this temporal development, beyond the overall *task time* (when the virtual ball has come to rest in the target area), that come to mind are marked in figure 3. These are

1. the number of changes of direction in the balls movement, i.e. cases where, in trying to stop inside the target, the ball is accelerated excessively against its momentary direction. In the example of figure 3 these points are marked by circles and numbered, the resulting index of their overall number (in one game or in average over a set of games) shall be called "*ball oscillations*".
2. Similarly, points in time where the direction of the applied acceleration changes, i.e. where the angle of the track changes sign (from positive to negative or vice versa, zero crossings of the red line) are numbered in the figure; the resulting index of the number of such occurring changes of sign is called "*inclination swaps*".
3. Another characteristic point in the ball trajectory is the moment when the ball reaches maximal velocity in one game. In figure 3 velocities are represented by black tangential lines.
4. Besides this *maximum velocity*, the time of its occurrence, "*max. velocity time*" and the ball position in this moment, "*max. velo. position*" are indices marked in the graph of figure 3 (by dotted lines) and discussed in the following.
5. Finally, one crucial moment from a perceptual side is the entering of the virtual ball into the target area. Derived indices are "*target reaching time*", "*target stopping time*", the time needed to stop the ball after having entered the target area (this is obviously the different between *task time* and *target reaching time*), " and the "*entry velocity*", i.e. the velocity with which the ball enters the target area.

3.2.1. Target reaching *and* stopping times

Table 4 gives a comparison of average *target reaching* and *stopping times* under the different conditions in the format known from table 3. It is seen that the average time between the ball entering the target area and it being finally stopped is in all cases shorter with sonic feedback than in the respective sets without sonic feedback (negative %-values only in the right, "*no*", columns of the lower chart of table 4). In accordance with the behaviour of *task times* (subsection 3.1) this effect is in the untrained set stronger with *rolling sound* (negative value in the left column of the lower chart of table 4) and vice versa stronger with *abstract sound* in the trained sets (positive %-value in the respective column).

[5] This is the common term for the role of the human interacting with a machine when the focus is on more low–level sensory–motor behaviour rather than higher cognitive aspects such as decision making.

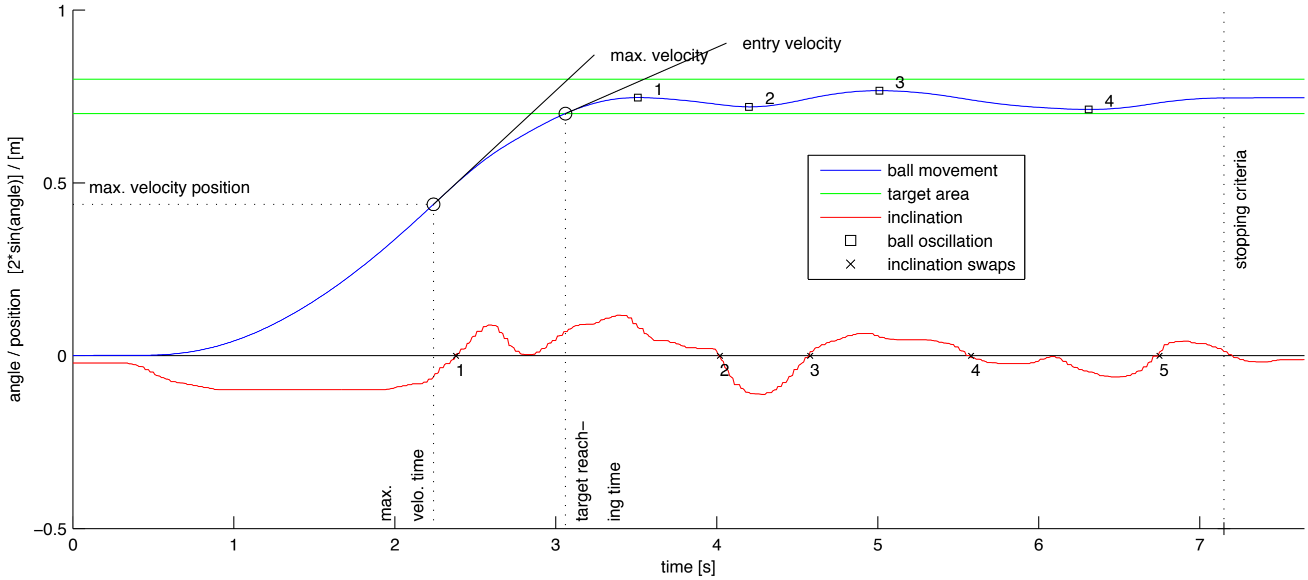

Figure 3: *Characteristic indices in the trajectories over time (x–axis in* s*) of the position of the virtual ball (blue, in* m*) and the underlying simultaneous development of the angle of the held control track in one game (the red curve depicts the sine of the angle, stretched by a factor of* 2 *for better visibility). As an effect of a simple model of "stick" and "roll" friction the virtual ball stays at rest near the start and end of the game also in short phases where the control track is not being held absolutely horizontal (not even completely still, at the resolution of the direction sensor).*

Relative differences of *target reaching times*, statistic significance	"untrained" *as*	"untrained" *no*	"trained" *as*	"trained" *no*
rs				
$\delta(\%)$	2.2	6.6	8	6.3
p	0.564	0.072	0.05	0.126
as				
$\delta(\%)$		4.3		−1.6
p		0.279		0.7

Relative differences of *target stopping times*, statistic significance	"untrained" *as*	"untrained" *no*	"trained" *as*	"trained" *no*
rs				
$\delta(\%)$	−18.6	−33.9	8.3	−11.4
p	0.086	0.001	0.451	0.273
as				
$\delta(\%)$		−18.8		−18.2
p		0.104		0.025

Table 4: *Relative differences of average* target reaching *and* stopping times *under the different conditions of training and sonic feedback*.

From the upper chart of table 4 it is noted that subjects in average took less time to reach the target area without sound than with, a remark that applies to all sets except *abstract sound* in the trained series where the *target reaching time* is slightly shorter with sound (only negative value in this chart). This faster reaching of the target is however *not* reflected by faster target completion, since we have seen the opposite being the case.

3.2.2. Maximum velocity

What has just been noted about *target reaching times* is basically found analogously for the *average maximum velocities* that the virtual ball reaches before entering the target: as seen in table 5 this velocity is in average always lower with sonic feedback (only negative %-values in the respective columns of the lower chart of table 5). The picture emerging from the last observations is that of a general tendency of subjects to overestimate the optimal maximum velocity that would form the best tradeoff in reaching the task fast but avoiding difficulties in subsequently stopping the virtual ball. At the same time the upper chart of table 5 shows that the average maximum velocity under both conditions of sonic feedback is significantly slower in the first "untrained" series while it shows much less adaptation with training in the "no sound" condition. Finally it is seen in table 5 that the maximum velocity is in average always slightly higher with the *abstract sound* than with *rolling sound*, a tendency that however in the trained series does not seem to conflict with the challenge of stopping the ball possibly efficiently and therefore to the contrary must be assumed to contribute to fast overall task performance with *abstract sound*.

Table 6 of the average positions at which the virtual ball reaches its maximum velocity shows that this *max. velo. position* is tendentially closer to the starting point of the trajectory with sonic feedback. This tendency, that is stronger in the trained series and here strongest for the *abstract sound*, can be seen as an index of

Average *maximum velocity*	*rs*	*as*	*no*
"untrained"	0.356	0.366	0.392
"trained"	0.386	0.4	0.407
$\delta(\%)$	8.2	9.4	3.9
p	0.049	0.05	0.358

Relative differences of *maximum velocities*, statistic significance	"untrained"		"trained"	
	as	*no*	*as*	*no*
rs				
$\delta(\%)$	-2.7	-9.1	-3.7	-5.3
p	0.5447	0.0157	0.3578	0.193
as				
$\delta(\%)$		-6.6		-1.6
p		0.1147		0.6999

Table 5: *Mean maximum velocities of the virtual ball under the different conditions of feedback and training.*

Average *max. velo. positions*	*rs*	*as*	*no*
"untrained"	0.36	0.37	0.37
"trained"	0.33	0.31	0.35
$\delta(\%)$	-9.3	-16.2	-6.3
p	0.033	0.000	0.178

Relative differences of *max. velo. positions*, statistic significance	"untrained"		"trained"	
	as	*no*	*as*	*no*
rs				
$\delta(\%)$	-2.8	-1.9	5.2	-5.1
p	0.5481	0.6792	0.2879	0.2609
as				
$\delta(\%)$		0.9		-9.8
p		0.8651		0.0228

Table 6: *Average distance from the starting point at which the virtual ball reaches its maximum velocity.*

more efficient acceleration behaviour with sound.

3.2.3. Entry velocity

The generally shorter *target stopping times* observed above (table 4) may seen in parallel with the average *entry velocities* that are tendentially lower with sonic feedback, as seen from according negative %-values in table 7. It is remarkable however that this effect is not present for *abstract sound* in the "trained" series (-0.7% value in the lower right of table 7). Apparently the stopping behaviour after training under this feedback condition is efficient enough to lead to significantly faster stopping than without sound also while the target area is here in average entered with the same velocity.

Relative differences of *entry velocities*, statistic significance	"untrained"		"trained"	
	as	*no*	*as*	*no*
rs				
$\delta(\%)$	-11.6	-17.2	-12.9	-13.5
p	0.065	0.005	0.034	0.05
as				
$\delta(\%)$		-6.4		-0.7
p		0.333		0.902

Table 7: *Average velocities of the virtual ball in the moment of entering the target area.*

3.2.4. Ball oscillations

Another characteristic of ball movement that comes to mind when considering the *target stopping times* is the number of oscillations, more precisely changes of direction, that the virtual ball undergoes in the course of being stopped inside the target. Table 8 shows the values of this index in the format used throughout this paper. It is

Average *ball oscillations*	*rs*	*as*	*no*
"untrained"	2.76	3.86	4.18
"trained"	2.57	2.52	3.18
$\delta(\%)$	-6.9	-34.8	-23.8
p	0.622	0.032	0.066

Relative differences of *ball oscillations*, statistic significance	"untrained"		"trained"	
	as	*no*	*as*	*no*
rs				
$\delta(\%)$	-28.5	-33.9	2	-19.4
p	0.0636	0.0019	0.909	0.2055
as				
$\delta(\%)$		-7.6		-20.9
p		0.635		0.1691

Table 8: *Average number of changes of direction of the ball's movement.*

seen that also this average value is tendentially smaller with sonic feedback. Probably most striking is the value in the untrained series under the *rolling sound* condition as the number of oscillations is here even smaller than without sound *after training*. This fact may be seen as an indication for a spontaneous understanding of this sonic feedback by subjects and, together with the corresponding values in the trained series, as a hint in the direction of the question posed in the introduction if sonic feedback may convey certain information, here velocity information, that is not equally well perceived through visual feedback.

3.2.5. Inclination swaps

Distinctive points in the control movement of the subject while performing the task of the experiment handled here are the moments where the direction of acceleration of the virtual ball changes. These represent "decisions" [6] of the controlling subject to actively accelerate the ball further in the direction of its current movement or to actively stop, i.e. accelerate against the momentary direction. It can be seen from the values in table 9 that the mean number of such *inclination swaps* is at both conditions of acoustic feedback smaller than without sound (only negative %-values in the right columns of table 9, lower chart). As was the

Average *inclination swaps*	*rs*	*as*	*no*
"untrained"	3.88	4.66	5.35
"trained"	3.96	3.73	4.1
$\delta(\%)$	1.9	−19.9	−23.4
p	0.820	0.008	0.001

Relative differences of *inclination swaps*, statistic significance				
	"untrained"		"trained"	
	as	*no*	*as*	*no*
rs				
$\delta(\%)$	−16.6	−27.4	6	−3.5
p	0.0404	0.0001	0.4441	0.6508
as				
$\delta(\%)$		−12.9		−8.9
p		0.1069		0.1356

Table 9: *Average number of changes of the direction of inclination*.

case with the *ball oscillations*, the situation for this index under *rolling sound* sticks out in that it is from the first series on smaller than without sound, even after training.

4. CONCLUSIONS

Summing up in general terms the results that were presented in some technical detail above the main conclusions shall be listed briefly.

- As in the preceding study (subsection 1.3) sonic feedback was found to lead to faster average performance of a task of target reaching on the basis of a balancing–metaphor. As a result of the design of the presented experiment there cannot remain any doubt that this performance improvement must be attributed to velocity information perceived from the sound.
- This effect is in the study found to be present also for the very good conditions of visual feedback on a wide–screen display spanning the whole range of the modelled scenario and the involved arm movements.
- The performance improvement is found to be, after training, larger with specially designed abstract sonic feedback than with the more ecological and more complex feedback of the *sound model* of rolling.
- The yet higher long–term performance achieved by an abstraction of the sonic behaviour in our case goes with a cost of increased necessary training. The work thus contributes to the general questions of intuitiveness and efficiency of ecologically based vs. rather abstract sounds in auditory display as proposed in the introduction.
- The analysis of movement trajectories by means of introduced characteristic indices (subsection 3.2) show various optimisations of control movements as a consequence of additional velocity information in sonic feedback, such as more efficient acceleration and stopping and a reduction of ball oscillations and changes of direction of inclination.
- The noted mechanisms must be assumed to happen widely unconscious so that the presented approach of measuring control movements under various feedback conditions can be regarded as complementing conventional techniques of psychoacoustic examination that rely on conscious reactions.

[6]The quotes are used to indicate that the question, if such an *inclination swap* is the result of conscious (as the word "decision" tendentially implies) or unconscious mechanisms, is here left open.

5. REFERENCES

[1] M. M. J. Houben, A. Kohlrausch, and D. J. Hermes, "Auditory cues determining the perception of size and speed of rolling balls," in *ICAD01*, Espoo, Finland, 2001, pp. 105–110.

[2] W. H. Warren and R. R. Verbrugge, "Auditory perception of breaking and bouncing events: a case study in ecological acoustics," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 10, no. 5, pp. 704–712, 1984.

[3] M. M. Blattner, D. A. Sumikawa, and R. M. Greenberg, "Earcons and icons: Their structure and common design principles," *Human Computer Interaction*, vol. 4, no. 1, pp. 11–44, 1989.

[4] W. W. Gaver, *Everyday listening and auditory icons*, Ph.D. thesis, University of California, San Diego, 1988.

[5] D. K. McGookin and S. A. Brewster, "Understanding concurrent earcons: Applying auditory scene analysis principles to concurrent earcon recognition," *ACM Transactions on Applied Perceptions*, vol. 1, no. 2, pp. 130–155, October 2004.

[6] "*The Sounding Object (SOb)*," European research project (IST-25287, < http//:www.soundobject.org >) as part of the *Disappearing Computer (DC)* proactive initiative (< http://www.disappearing-computer.org/ >).

[7] M. Rath, "An expressive real-time sound model of rolling.," in *Proceedings of the 6th "International Conference on Digital Audio Effects" (DAFx-03)*, London, United Kingdom, September 2003.

[8] "*InterSense*," < http://www.isense.com/ >.

[9] M. Rath and D. Rocchesso, "Informative sonic feedback for continuous human–machine interaction — controlling a sound model of a rolling ball," *IEEE Multimedia Special on Interactive Sonification*, pp. 60–69, April 2005.

[10] M.Rath, "Gestural exploitation of ecological information in continuous sonic feedback — the case of balancing a rolling ball," in *Proceedings of the 8th International Conference on Digital Audio Effects (DAFx'05)*, Madrid, Spain, 2005, pp. 98 – 103.

[11] Thomas B. Sheridan and William R. Ferrell, *Man–Machine Systems*, The MIT Press, Cambridge, Massachusetts, and London, England, 1974.

[12] Sukhan Lee and Hahk Sung Lee, "A kinesthetically coupled teleoperation: Its modelling and control," *IEEE*, 1991.